

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein -- protein search, using sw model

Run on: June 5, 2003, 16:12:50 ; Search time 27 Seconds  
 545.959 Million cell updates/sec (without alignments)

Title: US-09-869-414A-4  
 Perfect score: 2664  
 Sequence: 1 MAQALPWLLWMAGVLPAH.....CLRCURQQHDDADDISLK 501

Scoring table: BLOSUM62  
 Gapext 0.5  
 Scanned: 262574 seqs, 29422922 residues

1 number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%

Listing first 45 summaries

## Issued\_Patents\_AA:\*

```

1: /cn2_6/podata/1/iaa/5A_COMBO.pep:*
2: /cn2_6/podata/1/iaa/5B_COMBO.pep:*
3: /cn2_6/podata/1/iaa/6A_COMBO.pep:*
4: /cn2_6/podata/1/iaa/6B_COMBO.pep:*
5: /cn2_6/podata/1/iaa/PCUTS_COMBO.pep:*
6: /cn2_6/podata/1/iaa/backfiles1.pep:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	2656	99.7	501 4	US-09-548-372D-4	Sequence 4, Appli	
2	2656	99.7	501 4	US-09-548-367D-4	Sequence 4, Appli	
3	2650	99.5	501 4	US-09-191-2	Sequence 2, Appli	
4	2582	96.9	501 4	US-09-713-158-2	Sequence 2, Appli	
5	2557	96.4	501 4	US-09-548-372D-8	Sequence 8, Appli	
6	2506	96.4	501 4	US-09-1548-367D-8	Sequence 8, Appli	
7	2506	94.1	476 4	US-09-548-372D-6	Sequence 6, Appli	
8	2506	94.1	476 4	US-09-548-367D-6	Sequence 6, Appli	
9	2397	90.9	476 4	US-09-548-372D-7	Sequence 73, Appli	
10	2420	90.5	476 4	US-09-548-367D-7	Sequence 73, Appli	
11	2397	90.0	453 4	US-09-548-372D-30	Sequence 30, Appli	
12	2397	90.0	453 4	US-09-1548-367D-30	Sequence 30, Appli	
13	2288	90.0	459 4	US-09-548-372D-6	Sequence 32, Appli	
14	2397	90.0	459 4	US-09-548-367D-32	Sequence 32, Appli	
15	2315	86.9	774 4	US-09-191-4	Sequence 4, Appli	
16	2291	86.0	446 4	US-09-548-372D-22	Sequence 22, Appli	
17	2291	85.5	446 4	US-09-548-367D-22	Sequence 22, Appli	
18	2288	85.9	433 4	US-09-1548-372D-53	Sequence 26, Appli	
19	2288	85.9	433 4	US-09-548-367D-26	Sequence 26, Appli	
20	2288	85.9	459 4	US-09-548-372D-24	Sequence 24, Appli	
21	2288	85.9	459 4	US-09-548-367D-24	Sequence 24, Appli	
22	2247	84.4	428 4	US-09-548-372D-51	Sequence 51, Appli	
23	2247	84.4	428 4	US-09-548-367D-51	Sequence 51, Appli	
24	2247	84.4	434 4	US-09-1548-372D-53	Sequence 53, Appli	
25	2247	84.4	434 4	US-09-548-367D-53	Sequence 53, Appli	
26	2104	79.0	425 4	US-09-548-372D-28	Sequence 28, Appli	
27	2104	79.0	425 4	US-09-548-367D-28	Sequence 28, Appli	

## ALIGNMENTS

RESULT 1  
 US-09-548-372D-4  
 ; Sequence 4, Application US/09548372D  
 ; Patent No. 6420544  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GURNEY ET AL.  
 ; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE  
 ; FILE REFERENCE: 29915/62801  
 ; CURRENT APPLICATION NUMBER: US/09/548-372D  
 ; PRIORITY APPLICATION NUMBER: 2000-04-12  
 ; PRIORITY FILING DATE: 1999-05-23  
 ; PRIORITY APPLICATION NUMBER: US 09/404,133  
 ; PRIORITY FILING DATE: 1999-09-23  
 ; PRIORITY APPLICATION NUMBER: PCT/US99/20881  
 ; PRIORITY FILING DATE: 1999-05-23  
 ; PRIORITY APPLICATION NUMBER: US 60/101,594  
 ; PRIORITY FILING DATE: 1998-09-24  
 ; NUMBER OF SEQ ID NOS: 73  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 4  
 ; LENGTH: 501  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-548-372D-4

Query Match 99.7%; Score 2656; Pred. No. 2.9e-267;  
 Best Local Similarity 99.8%; Mismatches 0; Indels 0; Gaps 0;

Matches 500; Conservative 0;

1 MAQALPWLLWMAGVLPAHGTQHGRPLRSLGIGAPLGLRPRETDEPEEPGRGSF 60  
 1 MAQALPWLLWMAGVLPAHGTQHGRPLRSLGIGAPLGLRPRETDEPEEPGRGSF 60

Qy 1 YRDLEKGVVYVPTQCKWEGELGTDLVSIHPGPNTVRAIAITESDKFELLNGSMWEGIL 180  
 1 YRDLEKGVVYVPTQCKWEGELGTDLVSIHPGPNTVRAIAITESDKFELLNGSMWEGIL 180

Db 61 VEMVNLRGKSGOGYVEMTGSPTQTLNLTVDITGSSNFAVGAAPHPFLHRYTORQLSS 120  
 1 VEMVNLRGKSGOGYVEMTGSPTQTLNLTVDITGSSNFAVGAAPHPFLHRYTORQLSS 120

Qy 1 GLAYETIARPDDSLPFFDSLVKQTHVPNLFSLHLCGAGFLPNOSEVLSVGGSMIIGGI 240  
 1 GLAYETIARPDDSLPFFDSLVKQTHVPNLFSLHLCGAGFLPNOSEVLSVGGSMIIGGI 240

Db 121 YRDLEKGVVYVPTQCKWEGELGTDLVSIHPGPNTVRAIAITESDKFELLNGSMWEGIL 180  
 121 YRDLEKGVVYVPTQCKWEGELGTDLVSIHPGPNTVRAIAITESDKFELLNGSMWEGIL 180

Qy 181 GLAYETIARPDDSLPFFDSLVKQTHVPNLFSLHLCGAGFLPNOSEVLSVGGSMIIGGI 240  
 181 GLAYETIARPDDSLPFFDSLVKQTHVPNLFSLHLCGAGFLPNOSEVLSVGGSMIIGGI 240

Db 241 DHSLYTGSILWYTPIRREWYYEVIVRVEINGDLMKDCKEYNYDKSIVDSGTTNLRLPK 300  
 241 DHSLYTGSILWYTPIRREWYYEVIVRVEINGDLMKDCKEYNYDKSIVDSGTTNLRLPK 300

Db 241 DHSLYTGSILWYTPIRREWYYEVIVRVEINGDLMKDCKEYNYDKSIVDSGTTNLRLPK 300

QY 301 VFEAAVSKIAASSTEKFPGFWLGEOLVCWAGTTPWNIFPVISLYLMGEVTNQSFRIT 360  
 QY 301 VFEAAVSKIAASSTEKFPGFWLGEOLVCWAGTTPWNIFPVISLYLMGEVTNQSFRIT 360  
 Db 301 VFEAAVSKIAASSTEKFPGFWLGEOLVCWAGTTPWNIFPVISLYLMGEVTNQSFRIT 360  
 Qy 361 ILPQQYLPPVEDATSODCCYKFAISOSTSTGTWMGAVIMEGFYVVFDARKRIGFAVSAC 420  
 Db 361 ILPQQYLPPVEDATSODCCYKFAISOSTSTGTWMGAVIMEGFYVVFDARKRIGFAVSAC 420  
 Qy 361 ILPQQYLPPVEDATSODCCYKFAISOSTSTGTWMGAVIMEGFYVVFDARKRIGFAVSAC 420  
 Db 361 ILPQQYLPPVEDATSODCCYKFAISOSTSTGTWMGAVIMEGFYVVFDARKRIGFAVSAC 420  
 Qy 421 HVHDEERTAAVEGPFTLMDCGYNIPOTDESTMLTAYVMAICALFMLPLCLMVCQW 480  
 Db 421 HVHDEERTAAVEGPFTLMDCGYNIPOTDESTMLTAYVMAICALFMLPLCLMVCQW 480  
 Qy 421 HVHDEERTAAVEGPFTLMDCGYNIPOTDESTMLTAYVMAICALFMLPLCLMVCQW 480  
 Db 421 HVHDEERTAAVEGPFTLMDCGYNIPOTDESTMLTAYVMAICALFMLPLCLMVCQW 480

RESULT 3  
 QY 481 RCLRCLRQHDDFADDISLIK 501  
 Db 481 RCLRCLRQHDDFADDISLIK 501  
 QY 481 RCLRCLRQHDDFADDISLIK 501  
 Db 481 RCLRCLRQHDDFADDISLIK 501

RESULT 2  
 Sequence 4, Application US/09548367D  
 Patent No. 6440698  
 GENERAL INFORMATION:  
 APPLICANT: GURNEY ET AL.  
 TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES  
 TITLE OF INVENTION THEREOF  
 FILE REFERENCE: 2991576280H  
 CURRENT APPLICATION NUMBER: US/09/548.3 67D  
 CURRENT FILING DATE: 2000-04-12  
 PRIOR APPLICATION NUMBER: US 60/155,493  
 PRIOR FILING DATE: 1999-09-23  
 PRIOR APPLICATION NUMBER: US 09/404,133  
 PRIOR FILING DATE: 1999-09-23  
 PRIOR APPLICATION NUMBER: PCT/US99/20881  
 PRIOR FILING DATE: 1999-09-23  
 PRIOR APPLICATION NUMBER: US 60/101,594  
 PRIOR FILING DATE: 1998-09-24  
 NUMBER OF SEQ ID NOS: 73  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 4  
 LENGTH: 501  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-548-367D-4

Query Match 99.7%; Score 2056; DB 4; Length 501;  
 Best Local Similarity 98.8%; Pred. No. 2.9e-267;  
 Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 MAQALPWLLWMAGCYLPWLAHGTQHGIRLPURSIGGAPLGRLPRETDEEPEPGRGSF 60  
 1 MAQALPWLLWMAGCYLPWLAHGTQHGIRLPURSIGGAPLGRLPRETDEEPEPGRGSF 60

Db 61 VEMVDNLRGSGQGYVEMTGSPPTQLNLVDTGSSNFAVGAAAPHPLHRYYQYQRLSST 120  
 Qy 61 VEMVDNLRGSGQGYVEMTGSPPTQLNLVDTGSSNFAVGAAAPHPLHRYYQYQRLSST 120  
 Db 61 VEMVDNLRGSGQGYVEMTGSPPTQLNLVDTGSSNFAVGAAAPHPLHRYYQYQRLSST 120

Qy 121 YDLRKGVYVPTQGWEGTDLVSPHPNVTVRANIAATTESDEEFFINGSNWEGIL 180  
 Db 121 YDLRKGVYVPTQGWEGTDLVSPHPNVTVRANIAATTESDEEFFINGSNWEGIL 180

Qy 181 GLAYAETARPDSLEPFDSLQPKLQTHPVLNFSLHICGAGPPLNOSEVAISVGCSMIGI 240  
 Db 181 GLAYAETARPDSLEPFDSLQPKLQTHPVLNFSLHICGAGPPLNOSEVAISVGCSMIGI 240

Qy 1 MAQALPWLLWMAGCYLPWLAHGTQHGIRLPURSIGGAPLGRLPRETDEEPEPGRGSF 60  
 1 MAQALPWLLWMAGCYLPWLAHGTQHGIRLPURSIGGAPLGRLPRETDEEPEPGRGSF 60

Db 61 VEMVDNLRGSGQGYVEMTGSPPTQLNLVDTGSSNFAVGAAAPHPLHRYYQYQRLSST 120  
 Qy 61 VEMVDNLRGSGQGYVEMTGSPPTQLNLVDTGSSNFAVGAAAPHPLHRYYQYQRLSST 120

Db 61 VERDNLRGKSSQGYVEMTGSPPTQLNLVDTGSSNFAVGAAAPHPLHRYYQYQRLSST 120

Qy 121 YDLRKGVYVPTQGWEGTDLVSPHPNVTVRANIAATTESDKFTTNGSNWEGIL 180  
 Db 121 YDLRKGVYVPTQGWEGTDLVSPHPNVTVRANIAATTESDKFTTNGSNWEGIL 180





Page 5

Db	456	RCLRCRQHDDFADDISLLK	476
RESULT 8			
	US-09-548-367D-6		
	Sequence 6, Application US/09548367D		
	PATENT NO. 6440658		
	GENERAL INFORMATION:		
	APPLICANT: GURNEY ET AL.		
	TITLE OF INVENTION: ALZHEIMER'S DISEASE		
	FILE REFERENCE: 29915/6280H		
	CURRENT APPLICATION NUMBER: US/09/548		
	CURRENT FILING DATE: 2000-04-12		
	PRIOR APPLICATION NUMBER: US 60/1155 4		
	PRIOR FILING DATE: 1999-09-23		
	PRIOR APPLICATION NUMBER: US 09/404 1		
	PRIOR FILING DATE: 1999-09-23		
	PRIOR APPLICATION NUMBER: PCT/US99/20		
	PRIOR FILING DATE: 1999-09-23		
	PRIOR APPLICATION NUMBER: US 60/101,5		
	PRIOR FILING DATE: 1998-09-24		
	NUMBER OF SEQ ID NOS: 73		
	SEQ ID NO 6		
	LENGTH: 476		
	TYPE: PRT		
	ORGANISM: Homo sapiens		
	US-09-548-367D-6		
Qy		Query Match	94.1%
Db		Best Local Similarity	95.0%
Qy		Matches 476;	Score pred; Mis
Db		Conservative	0;
Qy	1	MAQALPWLMMAGYVLPAAHGTLQHGT	
Db	1	MAQALPWLMMAGYVLPAAHGTLQHGT	
Qy	61	YEMVDNLRGKSQQGYVEMTVGSPQQ	
Db	61	YEMVDNLRGKSQQGYVEMTVGSPQQ	
Qy	121	YRDLRKGVVYPYTOCKWEGELGTDLV	
Db	121	YRDLRKGVVYPYTOCKWEGELGTDLV	
Qy	181	GLAYAFIARPDDSLPEFFDSLVKQTH	
Db	181	GLAYAFIAR-----	
Qy	241	DHSLSLTSGLWYTPIRREWWYEVITVR	
Db	216	DHSLSLTSGLWYTPIRREWWYEVITVR	
Qy	301	VFEAAVKSIAASSTEKFPGFWLGEE	
Db	276	VFEAAVKSIAASSTEKFPGFWLGEE	
Qy	361	ILPQOYLRYVEDVATSQDDCYKFALS	
Db	336	ILPQOYLRYVEDVATSQDDCYKFALS	
Qy	421	HVHDEFRTAAVEGPFPVLDMEDCGYN	
Db	396	HVHDEFRTAAVEGPFPVLDMEDCGYN	
Qy	481	RCLRCRQHDDFADDISLLK 501	
Db	456	PCTPATPCQUDSEPDNTSITY 176	

US-09-548-372D-73	Sequence 73 , Application US/09548372D
; Patent No. 6420534	; GENERAL INFORMATION:
; ; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATE	; ; APPLICANT: GURNAY ET AL.
; ; FILE REFERENCE: 29915/62801	; ; CURRENT APPLICATION NUMBER: US/09/548,372D
; ; CURRENT FILING DATE: 2000-04-12	; ; PRIORITY NUMBER: US 60,155,493
; ; PRIORITY FILING DATE: 1999-09-23	; ; PRIORITY APPLICATION NUMBER: US 09/404,133
; ; PRIORITY FILING DATE: 1999-09-23	; ; PRIORITY APPLICATION NUMBER: PCT/US99/20881
; ; PRIORITY FILING DATE: 1999-09-23	; ; PRIORITY APPLICATION NUMBER: US 60,101,594
; ; PRIORITY FILING DATE: 1998-09-24	; ; PRIORITY FILING DATE: 1998-09-24
; ; NUMBER OF SEQ ID NOS: 73	; ; SOFTWARE: PatentIn version 3.1
; ; SEQ ID NO 73	; ; LENGTH: 476
; ; TYPE: PRT	; ; ORGANISM: Mus musculus
US-09-548-372D-73	
Query Match	90.9%
Best Local Similarity	91.8%
Matches	460;
Matches	Conservative
QY	1 MAQALPMLLNAGAVALPAHGTGIRPLRSIGGGAPGLRLRERETDEPEEE
Db	1 MAPALIWWLAWGSGMPLAQGTHGIRPLRSIGGGAPGLRLRERETDESEEE
QY	61 VEMVDNLRGKSGQQYYVEMTVGSPPOTLNILVDTGSSNFAVGAPHFPLHRYC
Db	61 VEMVDNLRGKSGQQYYVEMTVGSPPOTLNILVDTGSSNFAVGAPHFPLHRYC
QY	121 YRDLRKGVVYPVTQGWEGELGTDLVS1PHGPNTVVRANIAATESDKKFING
Db	121 YRDLRKGVVYPVTQGWEGELGTDLVS1PHGPNTVVRANIAATESDKKFING
Qy	181 GLAYAETIARPDDSLEPPFDLSVYKOTHVPNLFSLHLGAGFPNQSEVLSVGGG
Db	181 GLAYAETIAR-----LGAGFPNQSEVLSVGGG
Qy	241 DHSLYTGSWLYTPIRREMYYEIVLTVRVEINGQDLMQCKEYNYKSVIYDSGTT
Db	216 DHSLYTGSWLYTPIRREMYYEIVLTVRVEINGQDLMQCKEYNYKSVIYDSGTT
Qy	301 VFEAAVKSIKAASSTEKEPDGFNLGEQLVCWQAQGTPNNIFPV1SLYLMGEVTH
Db	276 VFEAAVKSIKAASSTEKEPDGFNLGEQLVCWQAQGTPNNIFPV1SLYLMGEVTH
Qy	361 ILPQOYLRPVEDATASODCYKFA1ISQSTGTWYAVIMEGFYVFDRAKRICK
Db	336 ILPQOYLRPVEDATASODCYKFAVSQSTGTWYAVIMEGFYVFDRAKRICK
Qy	421 HVHDEFRTAAVGPFTLMDGYNPTQDTESTMLTIAVMAICALFMLPLC
Db	396 HVHDEFRTAAVGPFTADMEDGYNPTQDTESTMLTIAVMAICALFMLPLC
Qy	481 RCLRCRQHDFAADDISLK 501
Db	456 RCLRCRQHDFAADDISLK 476

TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES  
 FILE REFERENCE: 29915/6280H  
 CURRENT APPLICATION NUMBER: US/09/548,367D  
 CURRENT FILING DATE: 2000-04-12  
 PRIOR APPLICATION NUMBER: US 60/155,493  
 PRIOR FILING DATE: 1999-09-23  
 PRIOR APPLICATION NUMBER: US/09/404,133  
 PRIOR FILING DATE: 1999-09-23  
 PRIOR APPLICATION NUMBER: PCT/US99/20881  
 PRIOR FILING DATE: 1999-09-23  
 PRIOR APPLICATION NUMBER: US 60/101,594  
 PRIOR FILING DATE: 1998-09-24  
 NUMBER OF SEQ ID NOS: 73  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO: 73  
 LENGTH: 476  
 TYPE: PRT  
 ORGANISM: Mus musculus

Query Match 90.9%; Score 2420 5; DB 4; Length 476;  
 Best Local Similarity 91.3%; Pred. No. 8e-243;  
 Matches 460; Conservative 5; Mismatches 11; Indels 25; Gaps 1;

QY 1 MAQALPWLWNGAGVLPAGTQHGLRPLSLGGAPLGLRPLTDEPEEPRGRGEF 60  
 DB 1 MAPALHWLWNGAGVLPAGTQHGLRPLSLGGAPLGLRPLTDEESEPRGRGEF 60  
 QY 61 VEMVDNLRKSGQQYVEMTGVSPPTLNLYDTGSNNFAVGAAAPHPLHRYQRLQST 120  
 DB 61 VEMVDNLRKSGQQYVEMTGVSPPTLNLYDTGSNNFAVGAAAPHPLHRYQRLQST 120  
 QY 121 YRDLRKGVVYPTQKGWEGLTDLYSTPHGPNTVRAIAITESDKFTTNSWEGIL 180  
 DB 121 YRDLRKGVVYPTQKGWEGLTDLYSTPHGPNTVRAIAITESDKFTTNSWEGIL 180  
 QY 181 GLAYAEIARPDDSLPEPFDLSLVKQTHPNFLSLHCGAGPLNQSEVLAvgGSMITGGI 240  
 DB 181 GLAYAEIARPDDSLPEPFDLSLVKQTHPNFLSLHCGAGPLNQSEVLAvgGSMITGGI 240  
 QY 241 DHSLYTGSLWTPTRREYYEVIVRVEINGDLKMDCKEKNYDKS1VDSGTNTNLPLKK 300  
 DB 241 DHSLYTGSLWTPTRREYYEVIVRVEINGDLKMDCKEKNYDKS1VDSGTNTNLPLKK 300  
 QY 301 VFEAAVKSIRAKASSTEKEPDPGEWLGEOLVCMQAGTTWNTEPVISLYLMGEVTNQSFRT 360  
 DB 301 VFEAAVKSIRAKASSTEKEPDPGEWLGEOLVCMQAGTTWNTEPVISLYLMGEVTNQSFRT 360  
 QY 361 ILPQQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVFDRARKRIGFVASAC 420  
 DB 361 ILPQQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVFDRARKRIGFVASAC 420  
 QY 421 HVHDEFRTAAVEGPVTLDMEDCNYNIPTODES 453  
 DB 421 HVHDEFRTAAVEGPVTLDMEDCNYNIPTODES 453  
 RESULT 11  
 US-09-548-372D-30  
 Sequence 30, Application US/09548367D  
 ; Sequence 30, Application US/09548367D  
 ; Patent No. 6420534  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GURNEY ET AL.  
 ; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES  
 ; FILE REFERENCE: 29915/6280H  
 ; CURRENT APPLICATION NUMBER: PCT/US99/20881  
 ; PRIOR FILING DATE: 1999-09-23  
 ; PRIOR APPLICATION NUMBER: US 60/101,594  
 ; PRIOR FILING DATE: 1998-09-24  
 ; CURRENT FILING DATE: 2000-04-12  
 ; CURRENT FILING DATE: 2000-04-12  
 ; NUMBER OF SEQ ID NOS: 73

PRIOR APPLICATION NUMBER: US 60/155,493  
 ; PRIOR FILING DATE: 1999-09-23  
 ; PRIOR APPLICATION NUMBER: US/09/404,133  
 ; PRIOR FILING DATE: 1999-09-23  
 ; PRIOR APPLICATION NUMBER: PCT/US99/20881  
 ; PRIOR FILING DATE: 1998-09-24  
 ; NUMBER OF SEQ ID NOS: 73

Query Match 90.0%; Score 2397; DB 4; Length 453;  
 Best Local Similarity 99.8%; Pred. No. 2e-240;  
 Matches 452; Conservative 0; Mismatches 0;  
 SEQ ID NO: 30  
 LENGTH: 453  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-548-372D-30

QY 1 MAQALPWLWNGAGVLPAGTQHGLRPLSLGGAPLGLRPLTDEPEEPRGRGEF 60  
 DB 1 MAQALPWLWNGAGVLPAGTQHGLRPLSLGGAPLGLRPLTDEPEEGRSSF 60  
 QY 61 VEMVDNLRKSGQQYVEMTGVSPPTLNLYDTGSNNFAVGAAAPHPLHRYQRLQST 120  
 DB 61 VEMVDNLRKSGQQYVEMTGVSPPTLNLYDTGSNNFAVGAAAPHPLHRYQRLQST 120  
 QY 121 YRDLRKGVVYPTQKGWEGLTDLYSTPHGPNTVRAIAITESDKFTTNSWEGIL 180  
 DB 121 YRDLRKGVVYPTQKGWEGLTDLYSTPHGPNTVRAIAITESDKFTTNSWEGIL 180  
 QY 181 GLAYAEIARPDDSLPEPFDLSLVKQTHPNFLSLHCGAGPLNQSEVLAvgGSMITGGI 240  
 DB 181 GLAYAEIARPDDSLPEPFDLSLVKQTHPNFLSLHCGAGPLNQSEVLAvgGSMITGGI 240  
 QY 241 DHSLYTGSLWTPTRREYYEVIVRVEINGDLKMDCKEKNYDKS1VDSGTNTNLPLKK 300  
 DB 241 DHSLYTGSLWTPTRREYYEVIVRVEINGDLKMDCKEKNYDKS1VDSGTNTNLPLKK 300  
 QY 301 VFEAAVKSIRAKASSTEKEPDPGEWLGEOLVCMQAGTTWNTEPVISLYLMGEVTNQSFRT 360  
 DB 301 VFEAAVKSIRAKASSTEKEPDPGEWLGEOLVCMQAGTTWNTEPVISLYLMGEVTNQSFRT 360  
 QY 361 ILPQQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVFDRARKRIGFVASAC 420  
 DB 361 ILPQQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVFDRARKRIGFVASAC 420  
 QY 421 HVHDEFRTAAVEGPVTLDMEDCNYNIPTODES 453  
 DB 421 HVHDEFRTAAVEGPVTLDMEDCNYNIPTODES 453  
 RESULT 12  
 US-09-548-367D-30  
 Sequence 30, Application US/09548367D  
 ; Sequence 30, Application US/09548367D  
 ; Patent No. 6440698  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GURNEY ET AL.  
 ; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES  
 ; FILE REFERENCE: 29915/6280H  
 ; CURRENT APPLICATION NUMBER: US/09/548,367D  
 ; PRIOR APPLICATION NUMBER: US 60/155,493  
 ; PRIOR FILING DATE: 1999-09-23  
 ; PRIOR APPLICATION NUMBER: US 09/404,133  
 ; PRIOR FILING DATE: 1999-09-23  
 ; PRIOR APPLICATION NUMBER: PCT/US99/20881  
 ; PRIOR FILING DATE: 1999-09-23  
 ; PRIOR APPLICATION NUMBER: US 60/101,594  
 ; PRIOR FILING DATE: 1998-09-24  
 ; NUMBER OF SEQ ID NOS: 73

SOFTWARE: PatentIn version 3.1  
; SEQ ID NO: 30  
; LENGTH: 453  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-548-367D-30

Query Match Similarity 90.0%; Score 2397; DB 4; Length 453;  
Best Local Similarity 99.8%; Pred. No. 2e-240; Indels 0; Gaps 0;

Matches 452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAQALPWLLWMGAGVLPAPAHTGTRPLRSLGGAPLGLRPRETDEPEPGRGSF 60  
Db 1 MAQALPWLLWMGAGVLPAPAHTGTRPLRSLGGAPLGLRPRETDEPEPGRGSF 60

Qy 1 MAQALPWLLWMGAGVLPAPAHTGTRPLRSLGGAPLGLRPRETDEPEPGRGSF 60  
Db 1 MAQALPWLLWMGAGVLPAPAHTGTRPLRSLGGAPLGLRPRETDEPEPGRGSF 60

Qy 61 VEMVDNLRGKSGQQYVEMTGVSPOTLNLYVDTGSSNFAVGAHPFLHRYQFQLSST 120  
Db 61 VEMVDNLRGKSGQQYVEMTGVSPOTLNLYVDTGSSNFAVGAHPFLHRYQFQLSST 120

Qy 121 YDLRKGVYVPTQKGWEGLTDVLSPHGPNTVRAAITESDKFPIINGSNWEGL 180  
Db 121 YDLRKGVYVPTQKGWEGLTDVLSPHGPNTVRAAITESDKFPIINGSNWEGL 180

Qy 181 GLAYAEIARPDSLEFFDSLVKQTHVPNLFSLQCGAGFPNQSEVLAvggsmiGGI 240  
Db 181 GLAYAEIARPDSLEFFDSLVKQTHVPNLFSLQCGAGFPNQSEVLAvggsmiGGI 240

121 YDLRKGVYVPTQKGWEGLTDVLSPHGPNTVRAAITESDKFPIINGSNWEGL 180  
121 YDLRKGVYVPTQKGWEGLTDVLSPHGPNTVRAAITESDKFPIINGSNWEGL 180

Qy 181 GLAYAEIARPDSLEFFDSLVKQTHVPNLFSLQCGAGFPNQSEVLAvggsmiGGI 240  
Db 181 GLAYAEIARPDSLEFFDSLVKQTHVPNLFSLQCGAGFPNQSEVLAvggsmiGGI 240

121 YDLRKGVYVPTQKGWEGLTDVLSPHGPNTVRAAITESDKFPIINGSNWEGL 180  
121 YDLRKGVYVPTQKGWEGLTDVLSPHGPNTVRAAITESDKFPIINGSNWEGL 180

Qy 301 VFEAAVSKIASSTEKFQDFWLGEOQVWQAGTPWNIFPVISLYLMSEVTNGSFRT 360  
Db 301 VFEAAVSKIASSTEKFQDFWLGEOQVWQAGTPWNIFPVISLYLMSEVTNGSFRT 360

Qy 361 ILPOQYLRPVEDVATSQDDCYKFAISOSSTGTVMGAVIMEGFYYVFDARKRIGFAVSAC 420  
Db 361 ILPOQYLRPVEDVATSQDDCYKFAISOSSTGTVMGAVIMEGFYYVFDARKRIGFAVSAC 420

Qy 361 ILPOQYLRPVEDVATSQDDCYKFAISOSSTGTVMGAVIMEGFYYVFDARKRIGFAVSAC 420  
Db 361 ILPOQYLRPVEDVATSQDDCYKFAISOSSTGTVMGAVIMEGFYYVFDARKRIGFAVSAC 420

Qy 301 VFEAAVSKIASSTEKFQDFWLGEOQVWQAGTPWNIFPVISLYLMSEVTNGSFRT 360  
Db 301 VFEAAVSKIASSTEKFQDFWLGEOQVWQAGTPWNIFPVISLYLMSEVTNGSFRT 360

Qy 361 ILPOQYLRPVEDVATSQDDCYKFAISOSSTGTVMGAVIMEGFYYVFDARKRIGFAVSAC 420  
Db 361 ILPOQYLRPVEDVATSQDDCYKFAISOSSTGTVMGAVIMEGFYYVFDARKRIGFAVSAC 420

Qy 421 HVHDEFRTAAVEGPFTLDMEDCGYNIPQDSES 453  
Db 421 HVHDEFRTAAVEGPFTLDMEDCGYNIPQDSES 453

RESULT 14  
US-09-548-367D-32  
; Sequence 32, Application US/09548367D  
; Patent No. 6440698

GENERAL INFORMATION:  
; APPLICANT: GURNEY ET AL.  
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE  
; FILE REFERENCE: 29915/6280H  
; CURRENT APPLICATION NUMBER: US/09/548,367D  
; CURRENT FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: US 60/155,493  
; PRIOR FILING DATE: 1999-09-23  
; PRIOR APPLICATION NUMBER: US 09/404,133  
; PRIOR FILING DATE: 1999-09-23  
; PRIOR APPLICATION NUMBER: PCT/US99/20881  
; PRIOR FILING DATE: 1999-09-23  
; PRIOR APPLICATION NUMBER: US 09/62801  
; CURRENT APPLICATION NUMBER: US/09/548,372D  
; CURRENT FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: US 60/155,493  
; PRIOR FILING DATE: 1999-09-23  
; NUMBER OF SEQ ID NOS: 73  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO: 32  
; LENGTH: 459  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-548-367D-32

Query Match Similarity 90.0%; Score 2397; DB 4; Length 459;  
Best Local Similarity 99.8%; Pred. No. 2.e-240; Indels 0; Gaps 0;

Qy 1 MAQALPWLLWMGAGVLPAPAHTGTRPLRSLGGAPLGLRPRETDEPEPGRGSF 60  
Db 1 MAQALPWLLWMGAGVLPAPAHTGTRPLRSLGGAPLGLRPRETDEPEPGRGSF 60

Qy 61 VEMVDNLRGKSGQQYVEMTGVSPOTLNLYVDTGSSNFAVGAHPFLHRYQFQLSST 120  
Db 61 VEMVDNLRGKSGQQYVEMTGVSPOTLNLYVDTGSSNFAVGAHPFLHRYQFQLSST 120

Qy 61 VEMVDNLRGKSGQQYVEMTGVSPOTLNLYVDTGSSNFAVGAHPFLHRYQFQLSST 120  
Db 61 VEMVDNLRGKSGQQYVEMTGVSPOTLNLYVDTGSSNFAVGAHPFLHRYQFQLSST 120

Query Match 86.9%; Score 2315; DB 4; Length 774;  
 Best Local Similarity 98.6%; Pred. No. 1.7e-231; Mismatches 0; Gaps 0;  
 Matches 436; Conservative 6;

---

Qy	121 YRDLEKGVVYVPYOGKWEGEGLTDVLSPHPNVTVRANIAITESDKF1NGSNWEGIL 180	Qy	58 GSFEVENDLNRGKSGQQYYEMTVCSPQTLNLYVDGSSNFAYGAAPHFLHRYQSQL 117
Db	121 YRDLEKGVVYVPYOGKWEGEGLTDVLSPHPNVTVRANIAITESDKF1NGSNWEGIL 180	Db	1 GSFEVENDLNRGKSGQQYYEMTVCSPQTLNLYVDGSSNFAYGAAPHFLHRYQSQL 60
Qy	181 GLAYEIAEPRDDSLPEPFDLVKOTHYNPLFSLHLCGAGFLINQESEVIAVGGSMLTGGI 240	Qy	118 SSTYRDLRKGVVYVPYOGKWEGEGLTDVLSPHPNVTVRANIAITESDKF1NGSNWEE 177
Db	181 GLAYEIAEPRDDSLPEPFDLVKOTHYNPLFSLHLCGAGFLINQESEVIAVGGSMLTGGI 240	Db	61 SSTYRDLRKGVVYEPYQKGWNEGELTDVLSPHPNVTVRANIAITESDKF1NGSNWEE 120
Qy	241 DHSLTGSLWTPIRREWWYEVILVVRVENGDLKMDCKEYNKDSVLSGTINRLPK 300	Qy	178 GTGLGLAYAEIARPDDSLPEPFDLVKOTHYNPLFSLHLCGAGEFLNQSEVIALASVGSMII 237
Db	241 DHSLTGSLWTPIRREWWYEVILVVRVENGDLKMDCKEYNKDSVLSGTINRLPK 300	Db	121 GTGLGLAYAEIARPDDSLPEPFDLVKOTHYNPLFSLHLCGAGEFLNQSEVIALASVGSMII 180
Qy	301 VFEAAVSKIRASSTEKFDPGFWLQEQLCWQAGTTPNWIFPVTSYLMGEVTNQSFIT 360	Db	238 GGDIDHSLYTGSLWTPIRREWWYEVILVVRVENGDLKMDCKEYNKDSVLSGTINRL 297
Db	301 VFEAAVSKIRASSTEKFDPGFWLQEQLCWQAGTTPNWIFPVTSYLMGEVTNQSFIT 360	Qy	181 GGDIDHSLYTGSLWTPIRREWWYEVILVVRVENGDLKMDCKEYNKDSVLSGTINRL 240
Qy	361 ILPQOYLRPVEDVATSQDCYKFAISQSSTGTYMGAVIMEGFYVVFDARKRIGFAVSAC 420	Qy	298 PKKVFEAAVSKIRASSTEKFDPGFWLQEQLCWQAGTTPNWIFPVTSYLMGEVTNQSF 357
Db	361 ILPQOYLRPVEDVATSQDCYKFAISQSSTGTYMGAVIMEGFYVVFDARKRIGFAVSAC 420	Db	241 PKKVFEAAVSKIRASSTEKFDPGFWLQEQLCWQAGTTPNWIFPVTSYLMGEVTNQSF 300
Qy	421 HVHDEFRTAAVEGPFPVTLMDMDCGYNIPQDTES 453	Qy	358 RITILPQOYLRPVEDVATSQDCYKFAISQSSTGTYMGAVIMEGFYVVFDARKRIGFAV 417
Db	421 HVHDEFRTAAVEGPFPVTLMDMDCGYNIPQDTES 453	Db	301 RITILPQOYLRPVEDVATSQDCYKFAISQSSTGTYMGAVIMEGFYVVFDARKRIGFAV 360

RESULT 15  
 US-09-009-191-4  
 ; Sequence 4, Application US/09009191  
 ; Patent No. 6313689

; GENERAL INFORMATION:  
 ; APPLICANT: POWELL, DAVID  
 ; APPLICANT: CHAPMAN, CONRAD  
 ; APPLICANT: MURPHY, KAY  
 ; APPLICANT: SMITH, TRUDI  
 ; TITLE OF INVENTION: ASP2  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: RAYER & PRESTIA  
 ; STREET: P.O. BOX 980  
 ; CITY: VALLEY FORGE  
 ; STATE: PA  
 ; COUNTRY: USA  
 ; ZIP: 19482

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/009,191  
 ; FILING DATE: 20-JAN-1998  
 ; CLASIFICATION:  
 ; PRIORITY APPLICATION DATA:  
 ; APPLICATION NUMBER: UK 9701684.4  
 ; FILING DATE: 28-JAN-1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: PRESTIA, PAUL F  
 ; REGISTRATION NUMBER: 23,031  
 ; REFERENCE/DOCKET NUMBER: GH-70368  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 610-407-0700  
 ; TELEFAX: 610-407-0701  
 ; TELEX: 846169  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 774 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein

US-09-009-191-4

Search completed: June 5, 2003, 16:18:45  
 Job time : 29 secs



Db 121 YRDLRKGVVPTQKGWEGELGTDLVSIPHGPNTVRANIAATESKFFINGSNWEGIL 180  
 Db 121 YRDLRKGVVPTQKGWEGELGTDLVSIPHGPNTVRANIAATESKFFINGSNWEGIL 180  
 Qy 181 GLAYAEIARPPDSLEPFDSLVKOTHVNLFSLHLCGAGFPINOSEVIAVGSMITGGI 240  
 Db 181 GLAYAEIARPPDSLEPFDSLVKOTHVNLFSLHLCGAGFPINOSEVIAVGSMITGGI 240  
 Db 181 GLAYAEIARPPDSLEPFDSLVKOTHVNLFSLHLCGAGFPINOSEVIAVGSMITGGI 240  
 Oy 241 DHSLYTGSWLYTPIREWWYEVILVVRVEINGDOLKMDCKEYNDKSIVDGTINRLPK 300  
 Db 241 DHSLYTGSWLYTPIREWWYEVILVVRVEINGDOLKMDCKEYNDKSIVDGTINRLPK 300  
 Qy 241 DHSLYTGSWLYTPIREWWYEVILVVRVEINGDOLKMDCKEYNDKSIVDGTINRLPK 300  
 Qy 301 VFEAAVSKSIASSTEKFDPGEOLCWOAGTTPNIFPVISLYLMEGYTNOSFRIT 360  
 Db 301 VFEAAVSKSIASSTEKFDPGEOLCWOAGTTPNIFPVISLYLMEGYTNOSFRIT 360  
 Qy 361 ILPQQYLRPVEDVATSDDCYKFAISQSSTGTVNGAVIMEGFYVVFDARKRIGFAVSAC 420  
 Db 361 ILPQQYLRPVEDVATSDDCYKFAISQSSTGTVNGAVIMEGFYVVFDARKRIGFAVSAC 420  
 Qy 421 HVHDEFRTAAVEGPFVTLMDCGYNIPQDESTLMITYVMAICALFMLPLCLMVCW 480  
 Db 421 HVHDEFRTAAVEGPFVTLMDCGYNIPQDESTLMITYVMAICALFMLPLCLMVCW 480  
 Qy 481 RCLRCRQHQDDFADDISLLK 501  
 Db 481 RCLRCRQHQDDFADDISLLK 501

## RESULT 2

US-09-794-927-4

; Sequence 4, Application US/09794927

;

Patent No.

US2001001632A1

;

GENERAL INFORMATION

;

APPLICANT:

Gurney, Mark E.

;

APPLICANT:

Bielkowski, Michael J.

;

APPLICANT:

Heinrikson, Robert L.

;

APPLICANT:

Parodi, Luis A.

;

APPLICANT:

Yan, Riqiang

;

TITLE OF INVENTION:

ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND

;

TITLE OF INVENTION:

USES

;

TITLE OF INVENTION:

THEREFOR

;

FILE REFERENCE:

28341/6280FG

;

CURRENT APPLICATION NUMBER:

US/09/794\_927

;

CURRENT FILING DATE:

2001-02-27

;

PRIOR APPLICATION NUMBER:

09/416\_901

;

PRIOR FILING DATE:

1999-10-13

;

PRIOR APPLICATION NUMBER:

60/155\_493

;

PRIOR APPLICATION NUMBER:

1999-09-23

;

PRIOR APPLICATION NUMBER:

09/404\_133

;

PRIOR APPLICATION NUMBER:

1999-09-23

;

PRIOR APPLICATION NUMBER:

PCT/US99/20881

;

PRIOR FILING DATE:

1998-09-24

;

PRIOR APPLICATION NUMBER:

60/101\_594

;

NUMBER OF SEQ ID NOS:

73

;

SOFTWARE:

PatentIn Ver. 2.0

;

SEQ ID NO 4

;

LENGTH: 501

;

TYPE: PRT

;

ORGANISM: Homo sapiens

;

US-09-794-927-4

## RESULT 3

US-09-795-847-4

; Sequence 4, Application US/09795847

; Patent No. US20010018208A1

; GENERAL INFORMATION

;

APPLICANT: Gurney, Mark E.

;

APPLICANT: Bielkowski, Michael J.

;

APPLICANT: Heinrikson, Robert L.

;

APPLICANT: Parodi, Luis A.

;

APPLICANT: Yan, Riqiang

;

TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND

;

TITLE OF INVENTION: USES

;

TITLE OF INVENTION: THESEFOR

;

FILE REFERENCE: 28341/6280FG

;

CURRENT APPLICATION NUMBER: US/09/795\_847

;

CURRENT FILING DATE: 2001-02-28

;

PRIOR APPLICATION NUMBER: 09/416\_901

;

PRIOR FILING DATE: 1999-10-13

;

PRIOR APPLICATION NUMBER: 60/155\_493

;

PRIOR FILING DATE: 1999-09-23

;

PRIOR APPLICATION NUMBER: 09/404\_133

;

PRIOR FILING DATE: 1999-09-23

;

PRIOR APPLICATION NUMBER: PCT/US99/20881

;

PRIOR FILING DATE: 1998-09-24

;

PRIOR APPLICATION NUMBER: 60/101\_594

;

NUMBER OF SEQ ID NOS: 73

;

SOFTWARE: PatentIn Ver. 2.0

;

SEQ ID NO 4

;

LENGTH: 501

;

TYPE: PRT

;

ORGANISM: Homo sapiens

;

US-09-795-847-4

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Gaps 0;

Indels 0; Gaps 0;

Matches 501; Conservative 0; Mismatches 0;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Best Local Similarity 100.0%; Pred. No. 5\_2e-241;

Matches 501; Conservative 0; Mismatches 0;

Query Match Score 2664; DB 10; Length 501;

Db 61 VEMVDNLRGKSGQGYVEMTVGSPQTNLNLDGSSNFAVGAAPHPFLHRYQRQLSST 120  
 Qy 121 YDLRKGVYPTYCKWEGELTDLVSIPHGPNTVRANIAATESDKFINGSNWEGL 180  
 Db 121 YDLRKGVYPTYCKWEGELTDLVSIPHGPNTVRANIAATESDKFINGSNWEGL 180  
 Qy 181 GLAYAEIARDDSLPFFDLVKOTHVNPFLSHLCAGFPNNOSEVLASVGGSMTIGGI 240  
 Db 181 GLAYAEIARDDSLPFFDLVKOTHVNPFLSHLCAGFPNNOSEVLASVGGSMTIGGI 240  
 Qy 241 DHSLYTGSLWYTPIREWWYEVILRVEINGDLKMDKEYNYDKSIVDSGTTNRLPKK 300  
 Db 241 DHSLYTGSLWYTPIREWWYEVILRVEINGDLKMDKEYNYDKSIVDSGTTNRLPKK 300  
 Qy 301 VFEAAVSKIAASSTEKFQDGFWLGEQLCWOAGTTPWNIFPVISLYLMGEVTNOSFRIT 360  
 Db 301 VFEAAVSKIAASSTEKFQDGFWLGEQLCWOAGTTPWNIFPVISLYLMGEVTNOSFRIT 360  
 Qy 361 ILPQQYLRPVEDVATSQDDCYKFAQSOSTGTVMGAVINEGFYVFDRAKRIGFAVSAC 420  
 Db 361 ILPQQYLRPVEDVATSQDDCYKFAQSOSTGTVMGAVINEGFYVFDRAKRIGFAVSAC 420  
 Qy 421 HVHDEFRTAAVEGPVTLMDCGYNIPQTDESTMLTIAYVMAAICALFPLCLMVCQN 480  
 Db 421 HVHDEFRTAAVEGPVTLMDCGYNIPQTDESTMLTIAYVMAAICALFPLCLMVCQN 480  
 Qy 481 RCLRCRQHDDFADDISLIK 501  
 Db 481 RCLRCRQHDDFADDISLIK 501  
 Qy 481 RCLRCRQHDDFADDISLIK 501  
 Db 481 RCLRCRQHDDFADDISLIK 501

## RESULT 4

US-09-794-743-4 ; Application US/09794743  
 ; Patent No. US20030191A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gurney, Mark E.  
 ; APPLICANT: Biernowski, Michael J.  
 ; APPLICANT: Heinrikson, Robert L.  
 ; APPLICANT: Parodi, Luis A.  
 ; APPLICANT: Yan, Riqiang  
 ; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES  
 ; TITLE OF INVENTION: THEREFOR  
 ; FILE REFERENCE: 28341/6280JL  
 ; CURRENT APPLICATION NUMBER: US/09/794,743  
 ; CURRENT FILING DATE: 2001-02-27  
 ; PRIOR APPLICATION NUMBER: 09/416,901  
 ; PRIOR FILING DATE: 1999-10-13  
 ; PRIOR APPLICATION NUMBER: 60/155,493  
 ; PRIOR FILING DATE: 1999-09-23  
 ; PRIOR APPLICATION NUMBER: 09/404,133  
 ; PRIOR FILING DATE: 1999-09-23  
 ; PRIOR APPLICATION NUMBER: PCT/US99/20881  
 ; PRIOR FILING DATE: 1999-09-23  
 ; PRIOR APPLICATION NUMBER: 60/101,594  
 ; PRIOR FILING DATE: 1998-09-24  
 ; NUMBER OF SEQ ID NOS: 73  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO: 4  
 ; LENGTH: 501  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-794-743-4

Query Match 100.0% Score 2664; DB 10; Length 501;  
 Best Local Similarity 100.0% Pred. No. 5.2e-241; Mismatches 0; Indels 0; Gaps 0;  
 Matches 501; Conservative 0; MisMatches 0; Del 0; Gaps 0;  
 Query 1 MAQALPWLLWMGAGVLPAGHPLRAHGTQHGIRLPLRSGLGGAPLGLRPRETDEEPEPGRGSF 60  
 Db 1 MAQALPWLLWMGAGVLPAGHPLRAHGTQHGIRLPLRSGLGGAPLGLRPRETDEEPEPGRGSF 60

61 VEMVDNLRGKSGQGYVEMTVGSPQTNLNLDGSSNFAVGAAPHPFLHRYQRQLSST 120  
 Db 61 VEMVDNLRGKSGQGYVEMTVGSPQTNLNLDGSSNFAVGAAPHPFLHRYQRQLSST 120  
 Qy 121 YDLRKGVYPTYCKWEGELTDLVSIPHGPNTVRANIAATESDKFINGSNWEGL 180  
 Db 121 YDLRKGVYPTYCKWEGELTDLVSIPHGPNTVRANIAATESDKFINGSNWEGL 180  
 Qy 181 GLAYAEIARDDSLPFFDLVKOTHVNPFLSHLCAGFPNNOSEVLASVGGSMTIGGI 240  
 Db 181 GLAYAEIARDDSLPFFDLVKOTHVNPFLSHLCAGFPNNOSEVLASVGGSMTIGGI 240  
 Qy 241 DHSLYTGSLWYTPIREWWYEVILRVEINGDLKMDKEYNYDKSIVDSGTTNRLPKK 300  
 Db 241 DHSLYTGSLWYTPIREWWYEVILRVEINGDLKMDKEYNYDKSIVDSGTTNRLPKK 300  
 Qy 301 VFEAAVSKIAASSTEKFQDGFWLGEQLCWOAGTTPWNIFPVISLYLMGEVTNOSFRIT 360  
 Db 301 VFEAAVSKIAASSTEKFQDGFWLGEQLCWOAGTTPWNIFPVISLYLMGEVTNOSFRIT 360  
 Qy 361 ILPQQYLRPVEDVATSQDDCYKFAQSOSTGTVMGAVINEGFYVFDRAKRIGFAVSAC 420  
 Db 361 ILPQQYLRPVEDVATSQDDCYKFAQSOSTGTVMGAVINEGFYVFDRAKRIGFAVSAC 420  
 Qy 421 HVHDEFRTAAVEGPVTLMDCGYNIPQTDESTMLTIAYVMAAICALFPLCLMVCQN 480  
 Db 421 HVHDEFRTAAVEGPVTLMDCGYNIPQTDESTMLTIAYVMAAICALFPLCLMVCQN 480  
 Qy 481 RCLRCRQHDDFADDISLIK 501  
 Db 481 RCLRCRQHDDFADDISLIK 501  
 Qy 481 RCLRCRQHDDFADDISLIK 501  
 Db 481 RCLRCRQHDDFADDISLIK 501

RESULT 5  
 US-09-794-748-4 ; Application US/09794748  
 ; Sequence 4, Application US/09794748  
 ; Patent No. US20030315A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gurney, Mark E.  
 ; APPLICANT: Biernowski, Michael J.  
 ; APPLICANT: Heinrikson, Robert L.  
 ; APPLICANT: Parodi, Luis A.  
 ; APPLICANT: Yan, Riqiang  
 ; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES  
 ; TITLE OF INVENTION: THEREFOR  
 ; FILE REFERENCE: 28341/6280JL  
 ; CURRENT APPLICATION NUMBER: US/09/794,748  
 ; CURRENT FILING DATE: 2001-02-27  
 ; PRIOR APPLICATION NUMBER: 09/416,901  
 ; PRIOR FILING DATE: 1999-10-13  
 ; PRIOR APPLICATION NUMBER: 60/155,493  
 ; PRIOR FILING DATE: 1999-09-23  
 ; PRIOR APPLICATION NUMBER: 09/404,133  
 ; PRIOR FILING DATE: 1999-09-23  
 ; PRIOR APPLICATION NUMBER: PCT/US99/20881  
 ; PRIOR FILING DATE: 1998-09-24  
 ; PRIOR APPLICATION NUMBER: 60/101,594  
 ; NUMBER OF SEQ ID NOS: 73  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO: 4  
 ; LENGTH: 501  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-794-748-4

Query Match 100.0% Score 2664; DB 10; Length 501;  
 Best Local Similarity 100.0% Pred. No. 5.2e-241; Matches 501; Conservative 0; MisMatches 0; Del 0; Gaps 0;  
 Query 1 MAQALPWLLWMGAGVLPAGHPLRAHGTQHGIRLPLRSGLGGAPLGLRPRETDEEPEPGRGSF 60

Db 1 MAQALPWLLWMGAGVLPAGHPLRAHGTQHGIRLPLRSGLGGAPLGLRPRETDEEPEPGRGSF 60

**RESULT 6**  
US-09-794-925-4  
*; Sequence 4, Application US/09794925  
; Patent No. US20020064819A1  
; GENERAL INFORMATION:  
; APPLICANT: Gurney, Mark E.  
; APPLICANT: Bienkowski, Michael J.  
; APPLICANT: Heinrikson, Robert L.  
; APPLICANT: Parodi, Luis A.  
; APPLICANT: Yan, Riqiang  
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES  
; TITLE REFERENCE: 28341/6280HG  
; CURRENT FILING DATE: 2001-02-27  
; PRIORITY APPLICATION NUMBER: 09/794,925  
; PRIORITY FILING DATE: 1999-09-23  
; PRIORITY FILING DATE: 1999-10-13  
; PRIORITY APPLICATION NUMBER: 60/155,493  
; PRIORITY FILING DATE: 1999-09-23  
; PRIORITY APPLICATION NUMBER: 09/404,133  
; PRIORITY FILING DATE: 1999-09-23  
; PRIORITY APPLICATION NUMBER: PC17US99/20881  
; PRIORITY FILING DATE: 1999-09-23  
; PRIORITY APPLICATION NUMBER: 60/101,594  
; PRIORITY FILING DATE: 1998-09-24  
; NUMBER OF SEQ ID NOS: 73  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 501  
; TYPE: PRT  
; ORGANISM: Homo sapiens*

**RESULT 7**  
US-09-681-442-4  
*; Sequence 4, Application US/09681442  
; Patent No. US20020081634A1  
; GENERAL INFORMATION:  
; APPLICANT: Gurney, Mark E.  
; APPLICANT: Bienkowski, Michael J.  
; APPLICANT: Heinrikson, Robert L.  
; APPLICANT: Parodi, Luis A.  
; APPLICANT: Yan, Riqiang  
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES  
; TITLE REFERENCE: 28341/6280HG  
; CURRENT FILING DATE: 2001-04-05  
; PRIORITY APPLICATION NUMBER: US/09/681,442  
; PRIORITY FILING DATE: 1999-10-13  
; PRIORITY APPLICATION NUMBER: 60/155,493  
; PRIORITY FILING DATE: 1999-09-23  
; PRIORITY APPLICATION NUMBER: 09/404,133  
; PRIORITY FILING DATE: 1999-09-23  
; PRIORITY APPLICATION NUMBER: PC17US99/20881  
; PRIORITY FILING DATE: 1999-09-23  
; PRIORITY APPLICATION NUMBER: 60/101,594  
; PRIORITY FILING DATE: 1998-09-24  
; NUMBER OF SEQ ID NOS: 73  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 501  
; TYPE: PRT  
; ORGANISM: Homo sapiens*

Query Match 100.0%; Score 2664; DB 10; Length 501;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-241; Indels 0; Gaps 0;  
 Matches 501; Conservative 0; Mismatches 0; Gaps 0;

Query Match 100.0%; Score 2664; DB 10; Length 501;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-241; Indels 0; Gaps 0;  
 Matches 501; Conservative 0; Mismatches 0; Gaps 0;

1 MAQALPWLWLMGAGVLPAGTQHGLPAHGTQHGLPLRSGGGATGLRLPRETDEEPEPGRGSF 60  
 1 MAQALPWLWLMGAGVLPAGTQHGLPAHGTQHGLPLRSGGGATGLRLPRETDEEPEPGRGSF 60  
 Db 61 VEMVDNLRGKSGQQYVEMTVGSPQTNLIVDTGSSNFAVGAAPHPHFLHRYYQRQLSST 120  
 Qy 61 VEMVDNLRGKSGQQYVEMTVGSPQTNLIVDTGSSNFAVGAAPHPHFLHRYYQRQLSST 120  
 Db 61 VEMVDNLRGKSGQQYVEMTVGSPQTNLIVDTGSSNFAVGAAPHPHFLHRYYQRQLSST 120  
 Qy 121 YDLRKGVVPTQYTKWEGELGTDLVSIHGPNTVRANIAALTESDKFINGSNWEGTL 180  
 Db 121 YDLRKGVVPTQYTKWEGELGTDLVSIHGPNTVRANIAALTESDKFINGSNWEGTL 180  
 Qy 181 GLAYAETARPDDSLPFFDSLVKOTHVPNLFSLHCGAEFLPNOSEVLAvgsmiGGI 240  
 Db 181 GLAYAETARPDDSLPFFDSLVKOTHVPNLFSLHCGAEFLPNOSEVLAvgsmiGGI 240  
 Db 181 GLAYAETARPDDSLPFFDSLVKOTHVPNLFSLHCGAEFLPNOSEVLAvgsmiGGI 240  
 Db 241 DHSLYTGSLWYTPIREWWYEVITYRVEEVITYRVEEVINGDLKMDCKEYNYDKSIVDGTNTNLRLPKK 300  
 Qy 241 DHSLYTGSLWYTPIREWWYEVITYRVEEVINGDLKMDCKEYNYDKSIVDGTNTNLRLPKK 300  
 Db 241 DHSLYTGSLWYTPIREWWYEVITYRVEEVINGDLKMDCKEYNYDKSIVDGTNTNLRLPKK 300  
 Db 301 VFEAAVSKIKASSTEKFPGFWLGEQJYQWQAGTPWNIFPVTSLYMGEVTNQSFRIT 360  
 Db 301 VFEAAVSKIKASSTEKFPGFWLGEQJYQWQAGTPWNIFPVTSLYMGEVTNQSFRIT 360  
 Db 361 ILPQOYLRPVEDVATSQDCYKFAISQSCTGMAVIMEGFYVFDRAKRIGFAVSAC 420  
 Qy 361 ILPQOYLRPVEDVATSQDCYKFAISQSCTGMAVIMEGFYVFDRAKRIGFAVSAC 420  
 Db 361 ILPQOYLRPVEDVATSQDCYKFAISOSSTGMAVIMEGFYVFDRAKRIGFAVSAC 420  
 Qy 421 HVDEFRTAAVEGFPFTLMDCDCYNTIPQTDESTLMTIAYMAICALFMPLCLMVQCW 480  
 Db 421 HVDEFRTAAVEGFPFTLMDCDCYNTIPQTDESTLMTIAYMAICALFMPLCLMVQCW 480  
 Qy 481 RCLRCRQHDDFADDISLK 501  
 Db 481 RCLRCRQHDDFADDISLK 501

Qy 1 MAQALPILLWAGVLPAAHGTQHGIPLRSLGLGAPLGLRLPRETDEEPEPGRGSF 60  
 Db 1 MAQALPILLWAGVLPAAHGTQHGIPLRSLGLGAPLGLRLPRETDEEPEPGRGSF 60

Qy 61 VEMVDNLRGKSGQQYVEMTYGSPPOTLNILVTGSSNFAVGAAPHFLHYQROLST 120  
 Db 61 VEMVDNLRGKSGQQYVEMTYGSPPOTLNILVTGSSNFAVGAAPHFLHYQROLST 120

Qy 121 YDLRKGVYVPTQKGMEGELGTDLYSIPHGPNTVRANIAITESDKFFINGSNWEGIL 180  
 Db 121 YDLRKGVYVPTQKGMEGELGTDLYSIPHGPNTVRANIAITESDKFFINGSNWEGIL 180

Qy 61 VEMVDNLRGKSGQQYVEMTYGSPPOTLNILVTGSSNFAVGAAPHFLHYQROLST 120  
 Db 61 VEMVDNLRGKSGQQYVEMTYGSPPOTLNILVTGSSNFAVGAAPHFLHYQROLST 120

Qy 121 YDLRKGVYVPTQKGWEGEGLDTLYSIPHGPNTVRANIAITESDKFFINGSNWEGIL 180  
 Db 121 YDLRKGVYVPTQKGWEGEGLDTLYSIPHGPNTVRANIAITESDKFFINGSNWEGIL 180

Qy 181 GLAYAEIARPDDSLPEFDLSLYKQTHYPNLFSLHLCGAGFPLNQSEVLASVGSMITGGI 240  
 Db 181 GLAYAEIARPDDSLPEFDLSLYKQTHYPNLFSLHLCGAGFPLNQSEVLASVGSMITGGI 240

Qy 181 GLAYAEIARPDDSLPEFDLSLYKQTHYPNLFSLHLCGAGFPLNQSEVLASVGSMITGGI 240  
 Db 181 GLAYAEIARPDDSLPEFDLSLYKQTHYPNLFSLHLCGAGFPLNQSEVLASVGSMITGGI 240

Qy 241 DHSLYTGSLLWTPIRREWWYEIVLVRVEINGDLKMDCKEKEYNDKSIVTSDSGTTNLRLPKK 300  
 Db 241 DHSLYTGSLLWTPIRREWWYEIVLVRVEINGDLKMDCKEKEYNDKSIVTSDSGTTNLRLPKK 300

Qy 301 VFEAAVSKIAASSTEKFDPDFWLGEOLVCNOAGTTPWNIEPVISLYLMGEVTNQSFRIT 360  
 Db 301 VFEAAVSKIAASSTEKFDPDFWLGEOLVCNOAGTTPWNIEPVISLYLMGEVTNQSFRIT 360

Qy 241 DHSLYTGSLLWTPIRREWWYEIVLVRVEINGDLKMDCKEKEYNDKSIVTSDSGTTNLRLPKK 300  
 Db 241 DHSLYTGSLLWTPIRREWWYEIVLVRVEINGDLKMDCKEKEYNDKSIVTSDSGTTNLRLPKK 300

Qy 181 GLAYAEIARPDDSLPEFDLSLYKQTHYPNLFSLHLCGAGFPLNQSEVLASVGSMITGGI 240  
 Db 181 GLAYAEIARPDDSLPEFDLSLYKQTHYPNLFSLHLCGAGFPLNQSEVLASVGSMITGGI 240

Qy 241 DHSLYTGSLLWTPIRREWWYEIVLVRVEINGDLKMDCKEKEYNDKSIVTSDSGTTNLRLPKK 300  
 Db 241 DHSLYTGSLLWTPIRREWWYEIVLVRVEINGDLKMDCKEKEYNDKSIVTSDSGTTNLRLPKK 300

Qy 301 VFEAAVSKIAASSTEKFDPDFWLGEOLVCNOAGTTPWNIEPVISLYLMGEVTNQSFRIT 360  
 Db 301 VFEAAVSKIAASSTEKFDPDFWLGEOLVCNOAGTTPWNIEPVISLYLMGEVTNQSFRIT 360

Qy 361 ILPQYLRPVEDVATSDDCYKFAISOSSTGTVMGAVIMEGEYVFDARKRIGFAVSAC 420  
 Db 361 ILPQYLRPVEDVATSDDCYKFAISOSSTGTVMGAVIMEGEYVFDARKRIGFAVSAC 420

Qy 301 VFEAAVSKIAASSTEKFDPDFWLGEOLVCNOAGTTPWNIEPVISLYLMGEVTNQSFRIT 360  
 Db 301 VFEAAVSKIAASSTEKFDPDFWLGEOLVCNOAGTTPWNIEPVISLYLMGEVTNQSFRIT 360

Qy 421 HVHDEFRTAAVEGFPVTLMEDCGYNIPQTDESTLMTIAYMAAICALFMPCLMVCQW 480  
 Db 421 HVHDEFRTAAVEGFPVTLMEDCGYNIPQTDESTLMTIAYMAAICALFMPCLMVCQW 480

Qy 361 ILPQYLRPVEDVATSDDCYKFAISOSSTGTVMGAVIMEGEYVFDARKRIGFAVSAC 420  
 Db 361 ILPQYLRPVEDVATSDDCYKFAISOSSTGTVMGAVIMEGEYVFDARKRIGFAVSAC 420

Qy 421 HVHDEFRTAAVEGFPVTLMEDCGYNIPQTDESTLMTIAYMAAICALFMPCLMVCQW 480  
 Db 421 HVHDEFRTAAVEGFPVTLMEDCGYNIPQTDESTLMTIAYMAAICALFMPCLMVCQW 480

RESULT 9  
 ; Sequence 104, Application US/10214932  
 ; Publication No. US20030100707A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HWANG, Inhwan  
 ; APPLICANT: KIM, Dae Heon  
 ; APPLICANT: LEE, Yong Jik  
 ; TITLE OF INVENTION: SYSTEM FOR DETECTING PROTEASE  
 ; FILE REFERENCE: ABP02/US  
 ; CURRENT APPLICATION NUMBER: US/10/214,932  
 ; CURRENT FILING DATE: 2002-08-08  
 ; NUMBER OF SEQ ID NOS: 133  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 104  
 ; LENGTH: 501  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-214-932-104

Query Match 99.7%; Score 2656; DB 9; Length 501;  
 Best Local Similarity 99.8%; Pred. No. 2.9e-240; Indels 0; Gaps 0;

Qy 1 MAQALPILLWAGVLPAAHGTQHGIPLRSLGLGAPLGLRLPRETDEEPEPGRGSF 60  
 Db 1 MAQALPILLWAGVLPAAHGTQHGIPLRSLGLGAPLGLRLPRETDEEPEPGRGSF 60

Qy 61 VEMVDNLRGKSGQQYVEMTYGSPPOTLNILVTGSSNFAVGAAPHFLHYQROLST 120  
 Db 61 VEMVDNLRGKSGQQYVEMTYGSPPOTLNILVTGSSNFAVGAAPHFLHYQROLST 120

Qy 121 YDLRKGVYVPTQKGWEGEGLDTLYSIPHGPNTVRANIAITESDKFFINGSNWEGIL 180  
 Db 121 YDLRKGVYVPTQKGWEGEGLDTLYSIPHGPNTVRANIAITESDKFFINGSNWEGIL 180

Qy 181 GLAYAEIARPDDSLPEFDLSLYKQTHYPNLFSLHLCGAGFPLNQSEVLASVGSMITGGI 240  
 Db 181 GLAYAEIARPDDSLPEFDLSLYKQTHYPNLFSLHLCGAGFPLNQSEVLASVGSMITGGI 240

Qy 241 DHSLYTGSLLWTPIRREWWYEIVLVRVEINGDLKMDCKEKEYNDKSIVTSDSGTTNLRLPKK 300  
 Db 241 DHSLYTGSLLWTPIRREWWYEIVLVRVEINGDLKMDCKEKEYNDKSIVTSDSGTTNLRLPKK 300

Qy 61 VEMVDNLRGKSGQQYVEMTYGSPPOTLNILVTGSSNFAVGAAPHFLHYQROLST 120  
 Db 61 VEMVDNLRGKSGQQYVEMTYGSPPOTLNILVTGSSNFAVGAAPHFLHYQROLST 120



Qy 134 QGKWEGEGLGTDLVSPHPNVTVRANIAITESDKFFINGSNWEGTLGLAYAEIARPDDS 193  
 Db 121 QGKWEGEGLGTDLVSPHPNVTVRANIAITESDKFFINGSNWEGTLGLAYAEIARPDDS 180  
 Qy 194 LEPFFDSVLVKOTHVPNLFLHLCGAGFPNQSEVLASVGSSMIGGIDHSLYTGSILWYTP 253  
 Db 181 LEPFFDSVLVKOTHVPNLFLHLCGAGFPNQSEVLASVGSSMIGGIDHSLYTGSILWYTP 240  
 Qy 254 IRREWWYEVIVRVETINGQDJKMDCKEYNNDKSTVDGGTNRLPKVFEAVKSIAAS 313  
 Db 241 IRREWWYEVIVRVETINGQDJKMDCKEYNNDKSTVDGGTNRLPKVFEAVKSIAAS 300  
 Qy 314 STEKFDPDGFWLGEQLVWCQAOACTTPWNIFPVTSLYLMGEYTQNOSFRITILPQQLRPVEDV 373  
 Db 301 STEKFDPDGFWLGEQLVWCQAOACTTPWNIFPVTSLYLMGEYTQNOSFRITILPQQLRPVEDV 360  
 Qy 374 ATSDQDCYKFAISQSSTGTGAVIMEGFYVVDARKRIGFAYSACHVIDEFRTAAVEG 433  
 Db 361 ATSDQDCYKFAISQSSTGTGAVIMEGFYVVDARKRIGFAYSACHVIDEFRTAAVEG 420  
 Qy 314 STEKFDPDGFWLGEQLVWCQAOACTTPWNIFPVTSLYLMGEYTQNOSFRITILPQQLRPVEDV 493  
 Db 301 STEKFDPDGFWLGEQLVWCQAOACTTPWNIFPVTSLYLMGEYTQNOSFRITILPQQLRPVEDV 480  
 Qy 374 ATSDQDCYKFAISQSSTGTGAVIMEGFYVVDARKRIGFAYSACHVIDEFRTAAVEG 433  
 Db 361 ATSDQDCYKFAISQSSTGTGAVIMEGFYVVDARKRIGFAYSACHVIDEFRTAAVEG 420  
 434 PFVTLMDMEDCGYNIPOTDESTLMTAYVMAICAFMLPLCLMVQWRCLRLRQHDDF 493  
 Db 421 PFVTLMDMEDCGYNIPOTDESTLMTAYVMAICAFMLPLCLMVQWRCLRLRQHDDF 480  
 Qy 494 ADDISLLK 501  
 Db 481 ADDISLLK 488

---

RESULT 13  
 US-09-796-264-2  
 ; Sequence 2, Application US/09796264  
 ; Patent No. US2002049303A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tang, Jordan J.N.  
 ; APPLICANT: Koelsch, Gerald  
 ; TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods  
 ; FILE REFERENCE: OMRF 179  
 ; CURRENT APPLICATION NUMBER: US/09/796,264  
 ; CURRENT FILING DATE: 2001-02-28  
 ; PRIOR APPLICATION NUMBER: 09/604,608  
 ; PRIOR FILING DATE: 2000-06-27  
 ; PRIOR APPLICATION NUMBER: 60/168,060  
 ; PRIOR FILING DATE: 1999-11-30  
 ; PRIOR APPLICATION NUMBER: 60/177,836  
 ; PRIOR FILING DATE: 2000-01-25  
 ; PRIOR APPLICATION NUMBER: 60/178,368  
 ; PRIOR FILING DATE: 2000-01-27  
 ; PRIOR APPLICATION NUMBER: 60/210,292  
 ; PRIOR FILING DATE: 2000-06-08  
 ; NUMBER OF SEQ ID NOS: 31  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO: 2  
 ; LENGTH: 488  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE: Purified Memapsin 2  
 ; OTHER INFORMATION: Putative peptide  
 ; OTHER INFORMATION: Amino Acids 28-48 are remnant putative peptide  
 ; OTHER INFORMATION: residues  
 ; OTHER INFORMATION: Amino Acids 58-61, 78, 80, 82-83, 116, 118-121,  
 ; OTHER INFORMATION: 156, 166, 174, 246, 274, 276, 278-281, 283, and  
 ; OTHER INFORMATION: residues in contact with the OM99-2  
 ; OTHER INFORMATION: inhibitor  
 ; OTHER INFORMATION: Amino acids 54-57, 61-68, 143-154, 165-168, 198-202, and  
 ; OTHER INFORMATION: 220-224 are N-lobe Beta Strands  
 ; OTHER INFORMATION: Amino Acids 184-191 and 210-217 are N-lobe Helices  
 ; OTHER INFORMATION: Amino acids 237-240, 247-249, 251-256, 259-260,  
 ; OTHER INFORMATION: 273-275, 282-285, 316-318, 331-336, 342-348,  
 ; OTHER INFORMATION: 354-357, 366-370, 372-375, 380-383, 390-395,  
 ; OTHER INFORMATION: 400-405, and 418-420 are C-lobe Beta Strands  
 ; OTHER INFORMATION: Amino Acids 286-299, 307-310, 350-353, 384-387,  
 ; OTHER INFORMATION: and 427-431 are C-lobe Helices  
 US-09-796-264-2

Query Match 96.9%; Score 2582; DB 9; Length 488;  
 Best Local Similarity 99.8%; Pred. No. 2.4e-233; Indels 0; Gaps 0;  
 Matches 487; Conservative 0; Mismatches -1;

Qy 14 AGVLPAHGHTQNGIRLPLRSGLGGAPLGLRPRETDEEEPFGRGSFVEMYDNLRGKSGQ 73  
 Db 1 AGVLPAHGHTQNGIRLPLRSGLGGAPLGLRPRETDEEEPFGRGSFVEMYDNLRGKSGQ 60

Qy 74 GYYVENTGSPQQTINLVDTGSSNFAGAAPPHPLHRYKQRQLSTYDLRKGVYVPT 133  
 Db 61 GYYVENTGSPQQTINLVDTGSSNFAGAAPPHPLHRYKQRQLSTYDLRKGVYVPT 120

Qy 134 QGKWEGEGLGTDLVSPHPNVTVRANIAITESDKFFINGSNWEGTLGLAYAEIARPDDS 193  
 Db 121 QGKWEGEGLGTDLVSPHPNVTVRANIAITESDKFFINGSNWEGTLGLAYAEIARPDDS 180

Qy 194 LEPEFFDSVLVKOTHVPNLFLHLCGAGFPNQSEVLASVGSSMIGGIDHSLYTGSILWYTP 253  
 Db 181 LEPEFFDSVLVKOTHVPNLFLHLCGAGFPNQSEVLASVGSSMIGGIDHSLYTGSILWYTP 240

Query Match 96.9%; Score 2582; DB 10; Length 488;  
 Best Local Similarity 99.8%; Pred. No. 2.4e-233;

Matches 487; Conservative 0; Mismatches 1; Indels 0; Gaps 0; OTHER INFORMATION: 376-377 are residues in contact with the OM99-2 ; OTHER INFORMATION: Inhibitor ; OTHER INFORMATION: Amino acids 54-57, 61-68, 73-80, 86-89, 109-111, 113-116, 122-134, 143-154, 165-168, 198-202, and 220-224 are N-lobe Beta Strands ; OTHER INFORMATION: Amino Acids 184-191 and 210-217 are N-lobe Helices ; OTHER INFORMATION: Amino Acids 184-191 and 210-217 are N-lobe Helices ; OTHER INFORMATION: Amino acids 237-240, 247-249, 251-256, 259-260, 273-275, 282-285, 316-318, 331-336, 342-348, 354-357, 366-370, 372-375, 380-383, 390-395, 400-405, and 418-420 are C-lobe Beta Strands ; OTHER INFORMATION: Amino Acids 286-299, 307-310, 350-353, 384-387, 396-399, 406-409, 427-431 are C-lobe Helices ; OTHER INFORMATION: and 427-431 are C-lobe Helices ; OTHER INFORMATION: and US-09-845-226-2

Query Match 96.9%; Score 2582; DB 10; Length 488;

Best Local Similarity 99.8%; Pred. No. 2.4e-233; 0; Gaps 0; Matches 487; Conservative 0; Indels 0; Gaps 0;

QY 14 AGVLAHGTGIRPLRSLGGAPLGLPREDEEPEPGRGSFYMVDNLRGSGQ 73  
Db 1 AGVLAHGTGIRPLRSLGGAPLGLPREDEEPEPGRGSFYMVDNLRGSGQ 60

QY 74 GYYVMTVGSPQTNLILYDTGSSNFAYGAAPHEFLHRYQROLSSTYRDLRGKVYPT 133  
Db 61 GYYVMTVGSPQTNLILYDTGSSNFAYGAAPHEFLHRYQROLSSTYRDLRGKVYPT 120

QY 134 QGWEGELGTDLVSPHPGPNVTVRNIAITESDKFFINGSNWEGILGJAYAETARPDS 193  
Db 121 QGWEGELGTDLVSPHPGPNVTVRNIAITESDKFFINGSNWEGILGJAYAETARPDS 180

QY 194 LEPEFFSLVKOTHVPLFSLHLCGAGFPNQSEVASYGSMIIGGIDSLYFGSLWTP 253  
Db 181 LEPEFFSLVKOTHVPLFSLHLCGAGFPNQSEVASYGSMIIGGIDSLYFGSLWTP 240

QY 254 IRREWYEVITVRYVEINGQDLKMDKEYNDKSIVDGTGNRLPKKFVFAAVSKIKAS 313  
Db 241 IRREWYEVITVRYVEINGQDLKMDKEYNDKSIVDGTGNRLPKKFVFAAVSKIKAS 300

QY 314 STEKEPDGFNLGEOLYCWQAOASTTPWNIFPVTSYLMEYTNOSEFRATLPQQYLRLPVEDY 373  
Db 301 STEKEPDGFNLGEOLYCWQAOASTTPWNIFPVTSYLMEYTNOSEFRATLPQQYLRLPVEDY 360

QY 374 ATSQDDCKFALQSOSSTGTVNGAVIMEGYVWEDRARKRIGFAYSACHVIDEFRTAAVEG 433  
Db 361 ATSQDDCKFALQSOSSTGTVNGAVIMEGYVWEDRARKRIGFAYSACHVIDEFRTAAVEG 420

QY 434 PFTVLDMDCGYNIPQTESTLMTIAYVMAACALEMPICLMLCWCRCRCLRQHDF 493  
Db 421 PFTVLDMDCGYNIPQTESTLMTIAYVMAACALEMPICLMLCWCRCRCLRQHDF 480

QY 494 ADDISLLK 501  
Db 481 ADDISLLK 488

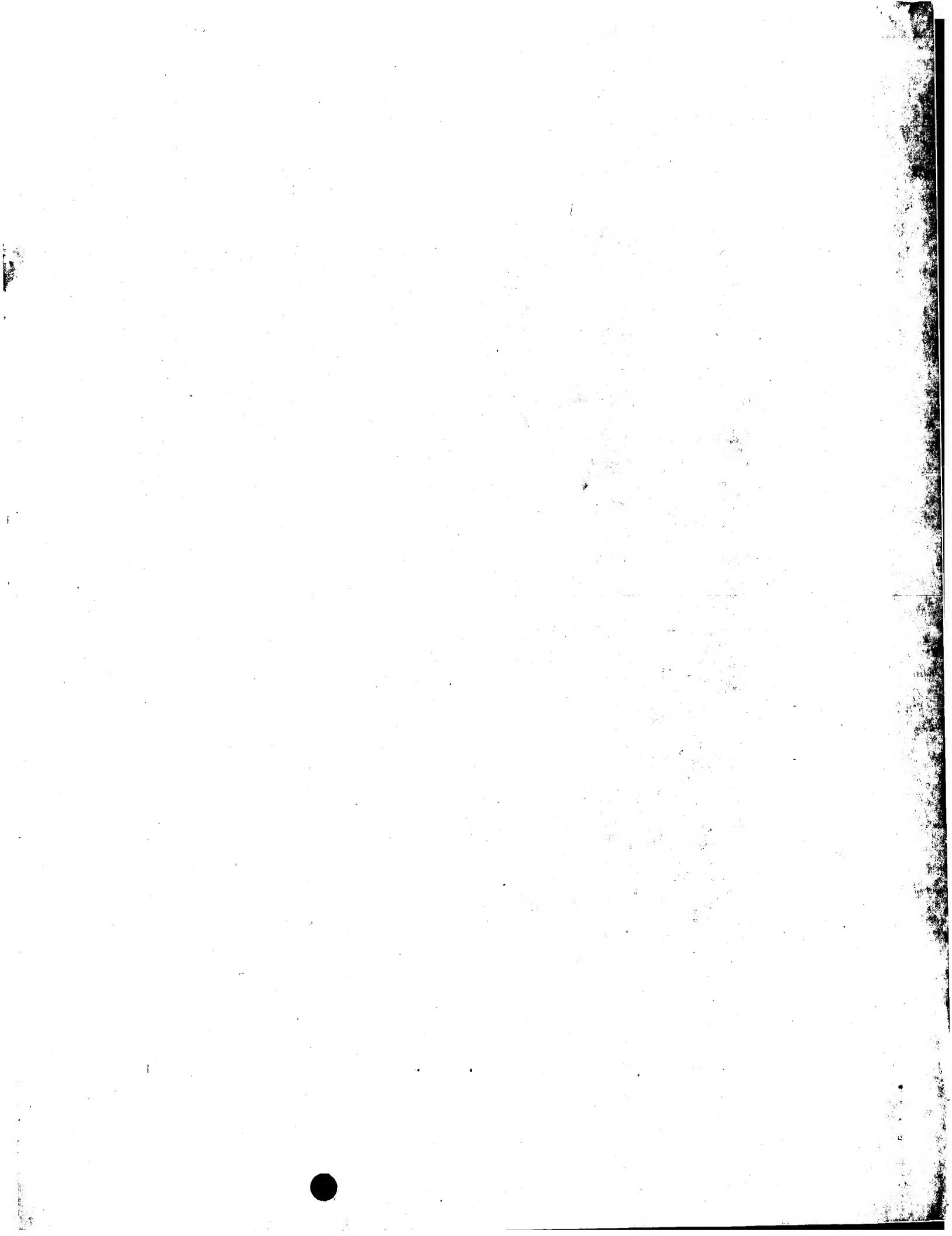
RESULT 14  
US-09-845-226-2  
Sequence 2, Application US/09845226  
; Patent No. US20020115600A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Jordan J.N.  
; APPLICANT: Hong, Lin  
; APPLICANT: Ghosh, Arun K.  
; TITLE OF INVENTION: Inhibitors of Memapsin 2 and Use Thereof  
; TITLE REFERENCE: OMRF 182  
; CURRENT APPLICATION NUMBER: US/09/845,226  
; CURRENT FILING DATE: 2001-04-30  
; PRIOR APPLICATION NUMBER: 09/7603,713  
; PRIOR FILING DATE: 2000-06-27  
; PRIOR APPLICATION NUMBER: 60/168,060  
; PRIOR FILING DATE: 1999-11-30  
; PRIOR APPLICATION NUMBER: 60/177,836  
; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 60/210,292  
; PRIOR FILING DATE: 2000-06-08  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO: 2  
; LENGTH: 488  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Purified Memapsin 2 are remnant putative propeptide  
; OTHER INFORMATION: Amino Acids 28-48 are residues in contact with the OM99-2  
; OTHER INFORMATION: residues 28-48 are remnant putative propeptide  
; OTHER INFORMATION: Amino Acids 58-61, 78, 80, 82-83, 116, 118-121, 156, 166, 174, 246, 274, 278-281, and  
; OTHER INFORMATION: and 427-431 are C-lobe Helices ; OTHER INFORMATION: and US-09-845-226-2

RESULT 15  
US-09-795-903A-3  
; Sequence 3, Application US/09795903A  
; Patent No. US2002016760A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Jordan J.N.  
; APPLICANT: Lin, Xini  
; APPLICANT: Koelsch, Gerald  
; TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods  
; TITLE OF INVENTION: Of Use Thereof  
; FILE REFERENCE: OMRF 179  
; CURRENT APPLICATION NUMBER: US/09/795,903A  
; CURRENT FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: 09/604,608  
; PRIOR FILING DATE: 2000-06-27  
; PRIOR APPLICATION NUMBER: 60/168,060  
; PRIOR FILING DATE: 1999-11-30  
; PRIOR APPLICATION NUMBER: 60/177,836

PRIOR FILING DATE: 2000-01-25  
 PRIOR APPLICATION NUMBER: 60/178,368  
 PRIOR FILING DATE: 2000-01-27  
 PRIOR APPLICATION NUMBER: 60/210,292  
 PRIOR FILING DATE: 2000-06-08  
 NUMBER OF SEQ ID NOS: 31  
 SOFTWARE: SeqMan Pro Version 6.0.2.1  
 SEQ ID NO: 3  
 LENGTH: 503  
 TYPE: PROT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 OTHER INFORMATION: Pro-memapsin 2  
 OTHER INFORMATION: Amino Acids 1-15 are vector-derived residues  
 OTHER INFORMATION: Amino Acids 16-64 are a putative pro peptide  
 OTHER INFORMATION: Amino Acids 1-13 are the T7 promoter  
 OTHER INFORMATION: Amino Acids 16-456 are Pro-memapsin 2-T1  
 OTHER INFORMATION: Amino Acids 16-421 are Pro-memapsin 2-T2  
 US-09-795-903A-3

	Very Match atches	Local Similarity at 87%; Conservative 0;	Score 96.9%	Pred. No. 2	Length 2582; DB 9; Length 503;	Indels 0;	Gaps 0;
Qy	14	AGVLAHGTOHQGIRPLRSLGGAPLGLRLPRETDEEPEPGRRGSFVEMVDNLRGKSGQ	73				
Db	16	AGVLAHGTOHQGIRPLRSLGGAPLGLRLPRETDEEPEPGRRGSFVEMVDNLRGKSGQ	75				
Qy	74	GYYVEMTYGSPPTQLNLVDTGSSNFAVGAAHPFLHRYYORQLSSTYRDLRGVYVPT	133				
Db	76	GYYVEMTYGSPPTQLNLVDTGSSNFAVGAAHPFLHRYYORQLSSTYRDLRGVYVPT	135				
Qy	134	QGKVEGEGLDTLVSIPHGENVTVRANIAITESDKFFINGSNVSEGILGLAYAEATRPDS	193				
Db	136	QGKVEGEGLDTLVSIPHGENVTVRANIAITESDKFFINGSNVSEGILGLAYAEATRPDS	195				
Qy	194	LEPFFDSLVKOTHYPVNLFSLHLCCAGFPNLNOSEVLASVGSMITGGIDHSLYTGSWYTP	253				
Db	196	LEPFFDSLVKOTHYPVNLFSLHLCCAGFPNLNOSEVLASVGSMITGGIDHSLYTGSWYTP	255				
Qy	254	IRREWWYEVITIVRVEINGODLKMDCKEYNDKSVTVDSCTTNLRLPKVFEAAVSKIAS	313				
Db	256	IRREWWYEVITIVRVEINGODLKMDCKEYNDKSVTVDSCTTNLRLPKVFEAAVSKIAS	315				
Qy	314	STERKPDGTWLGEOLVCAGTIPWNINIFVSLYLMGEVTNQSFRITIPQQVLRPVEDV	373				
Db	316	STERKPDGTWLGEOLVCAGTIPWNINIFVSLYLMGEVTNQSFRITIPQQVLRPVEDV	375				
Qy	374	ATSDODC7KFAISOSSTGTVMGAVTIMEGVYVFDARKRIGFAYSACHVDEFRTAAVEG	433				
Db	376	ATSDODC7KFAISOSSTGTVMGAVTIMEGVYVFDARKRIGFAYSACHVDEFRTAAVEG	435				
Qy	434	PFVTLMDMEDCGYNIPQTDESTLMTIAYVMAICALFMLPLCLMVCQWRCLRQHDDF	493				
Db	436	PFVTLMDMEDCGYNIPQTDESTLMTIAYVMAICALFMLPLCLMVCQWRCLRQHDDF	495				
Qy	494	ADDISLLK 501					
Db	496	ADDISLLK 503					

Search completed: June 5, 2003, 16:18:11  
 Job time : 25 secs



4 protein - protein search, using sw model			
Copyright (c) 1993 - 2003 Compugen Ltd.			
on:			June 5, 2003, 16:03:10 ; Search time 38 Seconds (without alignments)
on:			1756.804 Million cell updates/sec
title:			US-09-869-414A-4
perfect score:			2664
sequence:			1 MAQALPWLWMMGAVLPAH.....CLRCRLQQHDDFADDISLLK 501
scoring table:			BLOSUM62
Gapop 10.0 , Gapext 0.5			
Searched:			908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters:			908470
Amino DB seq length:			0
Maximum DB seq length:			2000000000
Post-processing:			Minimum Match 0% Maximum Match 100% Listing first 45 summaries
Database :			A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*			
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*			
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*			
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*			
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*			
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*			
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*			
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*			
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*			
10: /SIDS2/gcdat/genedata/geneseq/geneseq-emb1/AA1989.DAT:*			
11: /SIDS2/gcdat/genedata/geneseq/geneseq-emb1/AA1990.DAT:*			
12: /SIDS2/gcdat/genedata/geneseq/geneseq-emb1/AA1991.DAT:*			
13: /SIDS2/gcdat/genedata/geneseq/geneseq-emb1/AA1992.DAT:*			
14: /SIDS2/gcdat/genedata/geneseq/geneseq-emb1/AA1993.DAT:*			
15: /SIDS2/gcdat/genedata/geneseq/geneseq-emb1/AA1994.DAT:*			
16: /SIDS2/gcdat/genedata/geneseq/geneseq-emb1/AA1995.DAT:*			
17: /SIDS2/gcdat/genedata/geneseq/geneseq-emb1/AA1996.DAT:*			
18: /SIDS2/gcdat/genedata/geneseq/geneseq-emb1/AA1997.DAT:*			
19: /SIDS2/gcdat/genedata/geneseq/geneseq-emb1/AA1998.DAT:*			
20: /SIDS2/gcdat/genedata/geneseq/geneseq-emb1/AA1999.DAT:*			
21: /SIDS2/gcdat/genedata/geneseq/geneseq-emb1/AA2000.DAT:*			
22: /SIDS2/gcdat/genedata/geneseq/geneseq-emb1/AA2001.DAT:*			
23: /SIDS2/gcdat/genedata/geneseq/geneseq-emb1/AA2002.DAT:*			
RESULT 1 ID AAY88425 standard; Protein; 501 AA. XX AC AAY88425; XX DT 03-AUG-2000 (first entry) XX DE Human aspartyl protease 2 (a) (Asp2) amino acid sequence. XX KW Aspartyl protease; aspartate; amyloid precursor protein; APP; Asp 2; KW Alzheimer's disease; beta secretase site. XX PD 30-MAR-2000 XX PP 23-SEP-1999; 99WO-US20881. XX PR 24-SEP-1998; 98US-0101594. XX PA (PHAA ) PHARMACIA & UPJOHN CO. XX PI Gurney ME, Bienkowski MJ, Heinrikson RL, Parodi LA, Yan R; XX DR WPL; 2000-303209/26. DR N-PSDB; AAA15662. PT New enzyme designated human aspartase useful in research into the beta secretase site of cleaving amyloid protein precursor at the beta secretase site to produce amylvid. PT Alzheimer's Disease is capable of PT the beta secretase site to produce amylvid.			
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.			
SUMMARIES			
8			
result No.	Score	Match Length	DB ID
1	2664	100.0	501 21 AAY88425
2	2664	100.0	501 22 AAE10629
3	2664	100.0	501 22 AAE06859
4	2664	100.0	501 22 AAU06603
5	2664	100.0	501 22 AAU07202
6	2664	100.0	501 22 AAE02581
7	2664	100.0	501 23 ABT78590
8	2664	100.0	501 23 ABB06409
9	2856	99.7	501 21 AAY9776
10	2856	99.7	501 21 AAB07806

**Claim 48; Fig 2; 183pp; English.**

This sequence represents the human aspartyl protease 2 (Asp2) amino acid sequence. The invention relates to a protease (e.g. Asp<sub>2</sub>) capable of cleaving the beta secretase site of amyloid precursor protein (APP). The protease contains a sequence encoding a single amino acid sequence DTG and a sequence encoding DSG or DTG separated by 100-500 amino acids. When mutated, the APP gene causes an autosomal dominant form of Alzheimer's disease. APP localizes to the cell surface membrane and have a single C-terminal transmembrane domain. Proteolytic processing of APP produces the amyloid beta protein, which is possibly very important in Alzheimer's disease. The invention includes a nucleotide sequence encoding the protease, a vector containing the nucleotide sequence, and a cell line comprising the vector. Methods for screening for inhibitors of beta secretase activity are also given in the invention. The human aspartase protein and nucleotide sequences and the methods for identifying inhibitors of the protease, are useful in the treatment of and research into Alzheimer's disease.

Sequence 501 AA:

Query	Match	100.0%	Score 2664;	DB 21;	Length 501;
Best	Local Similarity	100.0%	Pred. No.	5.5e-263;	
Matches	Mismatches	0;	Indels	0;	Gaps 0;
Db	Conservative	0;			
Qy	1 MAGALPWLLWAGVDPFAHGTOHGRPLRPLSGLGAPLGLRIPRETDEPEPGRGSF 60				
Db	1 MAGALPWLLWAGVDPFAHGTOHGRPLRPLSGLGAPLGLRIPRETDEPEPGRGSF 60				
Qy	61 VEMYDNLRKSGGGYMMVTPSSPQTINLYVTGSSNFAVGAAPHPFLHRYYQRQLST 120				
Db	61 VEMYDNLRKSGGGYMMVTPSSPQTINLYVTGSSNFAVGAAPHPFLHRYYQRQLST 120				
Qy	121 YDLRKGVTVPTQKGELPDLYSPHPGNVTVTRANIAAITESDKKFITNSNWEGL 180				
Db	121 YDLRKGVTVPTQKGELPDLYSPHPGNVTVTRANIAAITESDKKFITNSNWEGL 180				
Qy	181 GLAYAEIARDDSLEPFDSLYKQTHVPNLFSLHCGAFPLNQSEVLAISGGSMIGI 240				
Db	181 GLAYAEIARDDSLEPFDSLYKQTHVPNLFSLHCGAFPLNQSEVLAISGGSMIGI 240				
Qy	241 DHSLYTGSLWYTPIRREWWYEVIVRVIEINGQDLKMDKEYNWDKSLIVDGTNTNLRLPKK 300				
Db	241 DHSLYTGSLWYTPIRREWWYEVIVRVIEINGQDLKMDKEYNWDKSLIVDGTNTNLRLPKK 300				
Qy	301 VEEAAVSKIAASSTEKEFPDFGEWLGEQLYCWOACTTPNNIFPVTSLYMGEVNQSPFRIT 360				
Db	301 VEEAAVSKIAASSTEKEFPDFGEWLGEQLYCWOACTTPNNIFPVTSLYMGEVNQSPFRIT 360				
Qy	361 ILPQQYLRPVEDVATSDDDCYKFAISQSSTGTYNGAVIMEGFYVVFDRARKRIGFAYSA 420				
Db	361 ILPQQYLRPVEDVATSDDDCYKFAISQSSTGTYNGAVIMEGFYVVFDRARKRIGFAYSA 420				
Qy	421 HVHDEFRTAAVEGPFTVLDMEDCGYNIPQDESTLMTAYVMAAICALFMLPLCLMYCQW 480				
Db	421 HVHDEFRTAAVEGPFTVLDMEDCGYNIPQDESTLMTAYVMAAICALFMLPLCLMYCQW 480				
Qy	481 RCLRCRQHDDFADDISLK 501				
Db	481 RCLRCRQHDDFADDISLK 501				

RESULT 2  
AAE10629 standard; Protein: 501 AA.

ID AAE10629  
XX  
AC AAE10629;  
XX  
DT 10-DEC-2001 (first entry)  
XX  
DE Human aspartyl protease 2(a) [hu-Asp2(a)] protein.

KW	Human; aspartyl protease 2(a); Asp2(a); amyloid precursor protein; APP; Alzheimer's disease; AD; dementia; neurofibrillary tangles; gliosis; amyloid plaque; neuronal loss; proteolytic; nootropic; neuroprotective; chromosome 11q23.3-24.1.				
XX					
PS	Homo sapiens.				
XX					
CC	Location/Qualifiers				
CC	Key 1..21				
CC	/label= Signal_peptide				
CC	Peptide 22..45				
CC	/label= Asp-2a-prepropeptide				
CC	FT 46..57				
CC	Peptide /label= Asp-2a-peptide				
CC	FT 58..501				
CC	Protein /label= Mature_human_Asp-2a_protein				
CC	FT 420..454				
CC	Region /label= Alpha-helical_spacer_region				
CC	Domain /label= Transmembrane_domain				
CC	FT 455..477				
CC	Domain /label= Cytoplasmic_domain				
CC	FT 478..501				
XX					
XX	GB2357767-A.				
XX	PN 04 -JUL-2001.				
XX	PR 22-SEP-1000; 20000GB-0023315.				
XX	PR 23-SEP-1999; 99US-0155493.				
XX	PR 23-SEP-1999; 99US-0404133.				
XX	PR 23-SEP-1999; 99WO-US-20881.				
XX	PR 13-OCT-1999; 99US-0416901.				
XX	PR 06-DEC-1999; 99US-0169232.				
XX	PA (PHAA ) PHARMACIA & UBJOHN CO.				
XX	PA Bierkowski MJ, Gurney M;				
XX	PA WPI: 2001-444208/48.				
XX	DR N-PSDB; AAD17865.				
XX	Polypeptide comprising fragments of human aspartyl protease with PT amyloid precursor protein processing activity and alpha secretase PT activity, for identifying modulators useful in treating Alzheimer's PT disease.				
XX	Example 2: Fig 2; 187pp; English.				
PS	The patent discloses human aspartyl protease 1 (hu-Asp1) or modified CC Asp1 proteins which lack transmembrane domain or amino terminal CC domain or cytoplasmic domain and retains alpha-secretase activity. The proteins CC and amyloid protein precursor (APP) processing activity. The proteins CC of the invention are useful for assaying hu-Asp1 alpha-secretase CC activity, which in turn is useful for identifying modulators of CC hu-Asp1 alpha-secretase activity, where modulators that increase CC hu-Asp1 alpha-secretase activity are useful for treating Alzheimer's CC disease (AD) which causes progressive dementia with consequent CC formation of amyloid plaques, neurofibrillary tangles, gliosis and CC neuronal loss. Hu-Asp1 protease substrate is useful for assaying CC hu-Asp1 proteolytic activity, by contacting hu-Asp1 protein with CC the substrate under acidic conditions and determining the level of CC hu-Asp1 proteolytic activity. The present sequence is long form of CC human Asp2 protein, designated as Asp2(a). Asp2 gene is localised CC on chromosome 11q23.3-24.1.				
XX	Sequence 501 AA:				
SQ	Query Match Score 100.0%; Pred. No. 5.5e-263; Length 501;				
XX	Best Local Similarity 100.0%; Pred. No. 5.5e-263;				
XX	Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1 MAQALPWLLWAGVDPFAHGTOHGRPLRPLSGLGAPLGLRIPRETDEPEPGRGSF 60				



CC	APP or its fragment containing an APP cleavage site recognizable at
CC	mammalian beta-secretase, and further comprising two lysine residues at the carboxyl terminus of the amino acid sequence of the mammalian APP or
CC	APP fragment. Also included in the invention are methods of identifying modulators and inhibitors of Asp2. Modulators and inhibitors of Asp2 are useful for treating Alzheimer's disease. APP is useful in methods for
CC	identifying inhibitors or modulators of human Asp2 activity and amyloid-beta (Abeta) peptide production. APP is also useful in designing therapeutics for the treatment or prevention of Alzheimer's disease.
CC	APP comprising the APP-Sw beta-secretase peptide sequence (NfDA), which is associated with increased levels of Abeta processing is useful in assays relating the Alzheimer's research. The expression vector is useful for recombinantly expressing APP. Nucleic acids that hybridise to APP oligonucleotides are useful as probes or primers. The probes are useful for detecting Hu-Asp nucleic acids in in vitro assays and in Northern and Southern blots. The present sequence is human Asp2(a).
XX	XX
Db	421 HVHDEFTAAVEGPFVTLDMEDCGYNIPQTDESTLMTAYVMAAICALEMLPLCLMVCW 480
Dy	481 RCLRCRQHDFADDISLKK · 501
Db	481 RCLRCRQHDFADDISLKK 501
RESULT 4	
ID	AAU06603 standard; protein: 501 AA.
XX	AAU06603;
XX	24-OCT-2001 (first entry)
XX	Human Aspartyl protease 2(a), Asp2(a).
XX	Human; Aspartyl protease; Asp2(a); beta secretase; nootropic; neuroprotective; amyloid protein precursor; APP; Alzheimer's disease; amyloid-beta; Abeta.
XX	Homo sapiens.
OS	Location/Qualifiers
FH	Key
Peptide	1..21 /label= Signal peptide
FT	22..45 /label= Pre_pro_peptide
Peptide	46..57 /label= Pro_peptide
FT	57..501 /label= Mature_Asp2(a)
Protein	420..454 /label= Alpha_helical_spacer_region
FT	455..477 /label= Transmembrane_domain
FT	478..501 /label= Cytoplasmic_domain
Region	Domian
FT	Domian
FT	WO20149098-A2.
PN	XX
XX	12-JUL-2001.
PD	XX
XX	09-MAY-2001; 2001WO-IB00798.
PF	XX
PA	09-MAY-2001; 2001WO-IB00798.
PA	(BIENKOWSKI M. J.
PA	(GURNY M. E.
PA	(HEINRICKSON R. L.
PA	(PARODI L. A.
PA	(YAN R.
XX	XX
PI	Bienkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Yan R;
XX	WPI; 2001-502549/55.
DR	N-PSDB; AAS11517.
XX	XX
PT	Novel purified polypeptide comprising fragment of mammalian aspartyl protease 2, lacking Asp2 transmembrane domain and retaining beta secretase activity of Asp2 useful for identifying inhibitors of Asp2 activity
PT	Claim 49; Fig 2; 185pp; English.
PS	XX
XX	The invention relates to a purified polypeptide comprising a fragment of mammalian aspartyl protease (Asp2) protein which lacks the Asp2 transmembrane domain and the Asp2 protein, and where the polypeptide and the fragment retain the beta-secretase activity of the mammalian Asp2 protein. The invention also details polynucleotides for the Asp2 protein and vectors expressing them, and a polypeptide (isoflote) of an amyloid protein precursor (APP) comprising the amino acid sequence of an
PT	RESULT 5
PT	AAU07202 ID AAU07202 standard; protein: 501 AA.
XX	XX
AC	AC
DT	24-OCT-2001 (first entry)
XX	Human aspartyl protease 2a (Asp-2a).
XX	XX
DE	DE
XX	XX
KW	Human; aspartyl protease 1; Asp-1; nootropic; neuroprotective;
KW	aspartyl protease 2; Asp2; amyloid protein precursor; APP;
CC	beta-secretase; Alzheimer's disease.
KW	..



Db	481	RCLRLRQQHDEADDISLKK	501
FTT	XX		
PN	W0200123533-A2.		
XX			
PD	05-APR-2001.		
XX			
PF	22-SEP-2000; 2000WO-US26080.		
XX			
PR	23-SEP-1999; 99US-0155493.		
XX			
PR	23-SEP-1999; 99WO-US20081.		
XX			
PR	13-OCT-1999; 99US-016901.		
XX			
PR	06-DEC-1999; 99US-0169232.		
XX			
(PHAA ) PHARMACIA & UPJOHN CO.			
PA	Gurney M, Bienkowski MJ;		
PPI	WPI; 2001-290516/30.		
XXX	N-PSDB; AAD06739.		
DR			
PT	Example 2; Fig 2; 189pp; English.		
XX			
PS	The present invention relates to enzymes for cleaving the alpha-secretase site of the amyloid precursor protein (APP) and methods of identifying those enzymes. The methods may be used to cleave the alpha-secretase cleavage site of the APP protein. The enzymes may be used to treat or modulate the progress of Alzheimer's disease. The present sequence is human aspartyl protease 2a (Asp 2a). Asp 2 has beta-secretase protease activity. Asp 2 gene is located on chromosome 11q23.3-24.1.		
XX			
CC	Sequence 501 <sup>a</sup> AA;		
CC	Query Match 100.0%; Score 2664; DB 22; Length 501;		
CC	Best Local Similarity 100.0%; Pred. No. 5.e-263; Mismatches 0; Indels 0; Gaps 0;		
CC	Matches 501; Conservative 0;		
Qy	1 MAQALPWLLWGMAGYLPAHGTOH GIRLPIRSGLGGAPLGLRPRETDEEPEPGRGSF 60		
Db	1 MAQALPWLLWGMAGYLPAHGTOH GIRLPIRSGLGGAPLGLRPRETDEEPEPGRGSF 60		
Qy	61 VEMVDNLRGKSGQYYVEMTVGSPOTLNIVDGTSSNFAVGAAHPHPELHRYFQROLST 120		
Db	61 VEMVDNLRGKSGQYYVEMTVGSPOTLNIVDGTSSNFAVGAAHPHPELHRYFQROLST 120		
Qy	121 YDLRKGVYYPTQKGWEGLTDLYSIPFVNTVRNATATESDKFPIFNSNWEGIL 180		
Db	121 YDLRKGVYYPTQKGWEGLTDLYSIPFVNTVRNATATESDKFPIFNSNWEGIL 180		
Qy	181 GLAYAETARPPDSLEPFDSLVKROTHVPLNLFSLHCGAGFPQLNOSEVLA SGGSMTGGI 240		
Db	181 GLAYAETARPPDSLEPFDSLVKROTHVPLNLFSLHCGAGFPQLNOSEVLA SGGSMTGGI 240		
Qy	301 VFEAVAKSIKAASSTERKEPDGFWLGEOLVCGAQATPWNITPVISLYLMGETVNOSFRIT 360		
Db	301 VFEAVAKSIKAASSTERKEPDGFWLGEOLVCGAQATPWNITPVISLYLMGETVNOSFRIT 360		
Qy	361 ILPOQYLRPVEDVATSDDCYKFAISQSSTGTMGAVIMEGFVYWDARKIGPAVSAC 420		
Db	361 ILPOQYLRPVEDVATSDDCYKFAISQSSTGTMGAVIMEGFVYWDARKIGPAVSAC 420		
Qy	421 HVHDEFRTAAVEGPFVTDMDMGYNIPTQDESTLMTIAYVMAAACALFMLPLCLMVCQW 480		
Db	421 HVHDEFRTAAVEGPFVTDMDMGYNIPTQDESTLMTIAYVMAAACALFMLPLCLMVCQW 480		
CC	Sequence 501 AA;		
CC	Query Match 100.0%; Score 2664; DB 23; Length 501;		
CC	Best Local Similarity 100.0%; Pred. No. 5.e-263;		

Qy	1 MAGALPWLWMMAGVLPAAHGTQHGRDPLRSLGGAPLGLRPRETDEEPBEPGRGGSF 60	PT modulators of beta secretase activity of aspartyl protease for treating Alzheimer's disease
Db	1 MAGALPWLWMMAGVLPAAHGTQHGRDPLRSLGGAPLGLRPRETDEEPBEPGRGGSF ,60	XX Claim 63: Page 118-119; 18pp; English.
Qy	61 VENVDNLRGKSGGGYVEMTVGSPQTTLNLYDTGSSNFAYGAAPHFLHRYQROLST 120	CC The present invention describes an isolated peptide (I) comprising a sequence of at least four amino acids, where the peptide is a substrate for conducting aspartyl protease assays. (I) has neuroprotective and nootropic activities, and can be used as an inhibitor of beta-secretase activity. A beta-secretase modulator from the present invention can be used for inhibiting beta-secretase activity in vivo, and in the manufacture of a medicament for the treatment of Alzheimer's disease.
Db	61 VERDNLRGKSGGGYVEMTVGSPQTTLNLYDTGSSNFAYGAAPHFLHRYQROLST 120	CC Pharmaceutical compositions from the present invention can be used for treating a disease or condition characterised by an abnormal beta-secretase activity. (I) is useful for identifying agents that modulate the activity of human Asp2 aspartyl protease (Hu-Asp2). (I) is useful as a core structure to construct derivatives. ABL4394 to ABL4925 and ABB0649 to ABB0593 represent sequences used in the exemplification of the present invention.
Qy	121 YDLRKGVYVPTQGKWEGEGLGTDLYSIPHGPNTVYRANIAAITESDKKFFINGSNVEGIL 180	CC
Db	121 YDLRKGVYVPTQGKNGEGLGTDLYSIPHGPNTVYRANIAAITESDKKFFINGSNVEGIL 180	CC
Qy	181 GLAYAETARPDSSLPEPFDLSVLYKOTHYNPNLFLHLCCAGFPLNQSEVLAvgGSMIGGI 240	CC
Db	181 GLAYAETARPDSSLPEPFDLSVLYKOTHYNPNLFLHLCCAGFPLNQSEVLAvgGSMIGGI 240	CC
Qy	241 DHSLYTGSLWYTPIRREMYEVTVRLVINGDQLKMDKEYNYDKS1TVDGTTNLRLPKK 300	CC
Db	241 DHSLYTGSLWYTPIRREMYEVTVRLVINGDQLKMDKEYNYDKS1TVDGTTNLRLPKK 300	CC
Qy	301 VFEAVKSIAASSTEKEPDGFMLGEOLVCMQAGTTPWNIFPVISLYLMGEVTNQSFRT 360	XX Sequence 501 AA;
Db	301 VFEAVKSIAASSTEKEPDGFMLGEOLVCMQAGTTPWNIFPVISLYLMGEVTNQSFRT 360	XX
Qy	361 ILPQQYLRPVEDATSDQDCYKFAISOSSTGTVGMGVIMEGYYVVFDRARRKGFAVSAC 420	Query Match 100.0%; Score 2664; DB 23; Length 501;
Db	361 ILPQQYLRPVEDATSDQDCYKFAISOSSTGTVGMGVIMEGYYVVFDRARRKGFAVSAC 420	Best Local Similarity 100.0%; Pred. No. 5.5e-263; Mismatches 0; Indels 0; Gaps 0;
Qy	421 HVHDEFRAAVEGPFVTLMDGYNIPQTDESTLMYTAYMMAICAFMLPLCLMVQW 480	Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	421 HVHDEFRAAVEGPFVTLMDGYNIPQTDESTLMYTAYMMAICAFMLPLCLMVQW 480	Qy 1 MAGALPWLWMMAGVLPAAHGTQHGRDPLRSLGGAPLGLRPRETDEEPBEPGRGGSF 60
Qy	421 HVHDEFRAAVEGPFVTLMDGYNIPQTDESTLMYTAYMMAICAFMLPLCLMVQW 480	Db 1 MAGALPWLWMMAGVLPAAHGTQHGRDPLRSLGGAPLGLRPRETDEEPBEPGRGGSF 60
Qy	421 HVHDEFRAAVEGPFVTLMDGYNIPQTDESTLMYTAYMMAICAFMLPLCLMVQW 480	Qy 61 VENDNLRGKSGGGYVEMTVGSPQTTLNLYDTGSSNFAYGAAPHFLHRYQROLST 120
Db	421 HVHDEFRAAVEGPFVTLMDGYNIPQTDESTLMYTAYMMAICAFMLPLCLMVQW 480	Db 61 VENDNLRGKSGGGYVEMTVGSPQTTLNLYDTGSSNFAYGAAPHFLHRYQROLST 120
Qy	481 RCLRCRQHDDFADDISLKK 501	Qy 61 VENDNLRGKSGGGYVEMTVGSPQTTLNLYDTGSSNFAYGAAPHFLHRYQROLST 120
Db	481 RCLRCRQHDDFADDISLKK 501	Db 61 VENDNLRGKSGGGYVEMTVGSPQTTLNLYDTGSSNFAYGAAPHFLHRYQROLST 120
Qy	481 RCLRCRQHDDFADDISLKK 501	Qy 121 YDRIRKGYYVPTQGKNGEGLGTDLYSIPHGPNTVYRANIAAITESDKKFFINGSNNEGIL 180
Db	481 RCLRCRQHDDFADDISLKK 501	Db 121 YDRIRKGYYVPTQGKNGEGLGTDLYSIPHGPNTVYRANIAAITESDKKFFINGSNNEGIL 180
RESULT 8		Qy 121 YDRIRKGYYVPTQGKNGEGLGTDLYSIPHGPNTVYRANIAAITESDKKFFINGSNNEGIL 180
ID	ABB06409 standard; Protein: 501 AA.	Db 121 YDRIRKGYYVPTQGKNGEGLGTDLYSIPHGPNTVYRANIAAITESDKKFFINGSNNEGIL 180
XX		Qy 181 GLAYAETARPDSSLPEPFDLSVLYKOTHYNPNLFLHLCCAGFPLNQSEVLAvgGSMIGGI 240
AC		Db 181 GLAYAETARPDSSLPEPFDLSVLYKOTHYNPNLFLHLCCAGFPLNQSEVLAvgGSMIGGI 240
XX		Qy 241 DHSEYTGSLWYTPIRREMYEVTVRLVINGDQLKMDKEYNYDKS1TVDGTTNLRLPKK 300
DT	31-MAY-2002 (first entry)	Db 241 DHSEYTGSLWYTPIRREMYEVTVRLVINGDQLKMDKEYNYDKS1TVDGTTNLRLPKK 300
XX	Human aspartyl protease protein sequence SEQ ID NO:2.	Qy 301 VFEAVKSIAASSTEKEPDGFMLGEOLVCMQAGTTPWNIFPVISLYLMGEVTNQSFRT 360
DE		Db 301 VFEAVKSIAASSTEKEPDGFMLGEOLVCMQAGTTPWNIFPVISLYLMGEVTNQSFRT 360
XX	Beta-secretase; enzyme; cleavage site; amyloid protein precursor; APP; aspartyl protease; neuroprotective; nootropic; beta-secretase inhibitor; Alzheimer's disease.	Qy 361 ILPQQYLRPVEDATSDQDCYKFAISOSSTGTVGMGVIMEGYYVVEDARKRIGFAYSAC 420
XX	Homo sapiens.	Db 361 ILPQQYLRPVEDATSDQDCYKFAISOSSTGTVGMGVIMEGYYVVEDARKRIGFAYSAC 420
OS		Qy 421 HVHDEFRAAVEGPFVTLMDGYNIPQTDESTLMYTAYMMAICAFMLPLCLMVQW 480
PN	WO200206306-A2.	Db 421 HVHDEFRAAVEGPFVTLMDGYNIPQTDESTLMYTAYMMAICAFMLPLCLMVQW 480
XX		Qy 481 RCLRCRQHDDFADDISLKK 501
PD	24-JAN-2002.	Db 481 RCLRCRQHDDFADDISLKK 501
XX	19-JUL-2001; 2001WO-US23035.	XX AC AAY94767;
PF		XX XX 12-FEB-2001 (first entry)
XX	19-JUL-2000; 2000US-219795P.	XX DE Human beta-secretase amino acid sequence.
PR	12-MAR-2001; 2001US-275251P.	XX PT Novel substrates for human aspartyl protease useful for identifying

KW Beta-secretase; enzyme; amyloid plaque; Alzheimer's disease; human;  
Down's syndrome; amyloid angiopathy; gene therapy; neuroprotective.  
XX

OS Homo sapiens.

FH Location/Qualifiers  
1.45  
/label= putative signal peptide  
46..501  
/label= Beta-secretase

FT Protein

XX WO200058479-A1.

PD 05-OCT-2000.

XX 23-MAR-2000; 2000WO-US07755.

XX 26-MAR-1999; 99US-0277229.

PR (AMGEN) AMGEN INC.

PI Citron M, Vassar RJ, Bennett BD;

XX WPI; 2000-594643/56.

DR N-PSDB; AAA28278.

XX Isolated beta-secretase nucleic acids and encoded polypeptides, useful  
for diagnosis and gene therapy of Alzheimer's disease -

PT XX

XX P1; Fig 4; 145pp; English.

CC This invention relates to 3 nucleotide sequences encoding beta-secretase  
proteins. Beta-secretase is an enzyme involved in the production of one  
of the components of amyloid plaques involved in Alzheimer's disease. The  
invention includes an expression vector comprising the nucleotide  
sequence, a host cell comprising the expression vector, and a process for  
producing the protein through culturing the transformed cells. Also  
included in the invention are a polypeptide derivative of the  
beta-secretase protein, a fusion protein comprising beta-secretase fused  
to a heterologous amino acid sequence, and a method for modulating the  
levels of beta-secretase polypeptide in a mammal comprising administering  
the polynucleotide sequence. Beta-secretase exhibits neuroprotective and  
nootropic activity. The beta-secretase nucleotide sequence may be used to  
map locations of the beta-secretase gene and related genes on chromosomes  
and as hybridization probes in diagnostic assays to test for the presence  
of beta-secretase DNA or RNA, such as in Alzheimer's disease, Down's  
syndrome, and amyloid angiopathy. The nucleotide sequence may also be  
used as anti-sense inhibitors of beta-secretase expression, in gene  
therapy of Alzheimer's disease, and for the identification of compounds  
that modulate beta-secretase activity. Antibodies to the beta-secretase  
protein may be used for in vitro and in vivo diagnostic purposes to  
detect the presence of beta-secretase polypeptide in a body fluid or cell  
sample. The present sequence represents the human beta-secretase protein.

CC XX

SQ Sequence 501 AA;

Query Match 99.7%; Score 2656; DB 21; Length 501;  
Best Local Similarity 99.8%; Pred. No. 3..6e-221;  
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAQALPWLWAGVLPAGLPAHGTOHGRILPLRSGLGGPLRPRETDEEPEPGRRGSF 60  
1 MAQALPWLWAGVLPAGLPAHGTOHGRILPLRSGLGGPLRPRETDEEPEPGRRGSF 60

DB 61 VEMVDNLRGKSGQGYVENTVGSPQTNLVDTGSNFAYGAAPHEFLHRYYQRQLSS'R'120  
1 VEMVDNLRGKSGQGYVENTVGSPQTNLVDTGSNFAYGAAPHEFLHRYYQRQLSS'120

QY 121 YRDLRKGVYYPTQCKWEGFLGTLSIPHGPNTVTNIAATTESKEFFINGSNWEGIL 180  
121 YRDLRKGVYYPTQCKWEGFLGTLSIPHGPNTVTNIAATTESKEFFINGSNWEGIL 180

DB 181 GLAYAEIARPDDSLFPFDLSIVKOTHVNLFSLHLCGAGFPLNOSEYLASYGGSMLIGGI 240

QY 181 GLAYAEIARPDDSLFPFDLSIVKOTHVNLFSLHLCGAGFPLNOSEYLASYGGSMLIGGI 240

Db 241 DHSLYTGSLWYTPIRREYYEVITIVRVEINQDILKADCKKEYNYDSIVDGTGTTNRLPKK 300

QY 241 DHSLYTGSLWYTPIRREYYEVITIVRVEINQDILKADCKKEYNYDSIVDGTGTTNRLPKK 300

Db 301 VFEAAVKSIIASSTEKRPDPGFWLGEOLVCMRAGTIPWNIFPVISIYLMGEVTNQSFRIT 360

QY 301 VFEAAVKSIIASSTEKRPDPGFWLGEOLVCMRAGTIPWNIFPVISIYLMGEVTNQSFRIT 360

Db 361 ILPOQYLRPVEDYATSDDDCYKFAISQSSTGTMGAVIMEGFYVVFDRARRKIGFAVSAC 420

QY 361 ILPOQYLRPVEDYATSDDDCYKFAISQSSTGTMGAVIMEGFYVVFDRARRKIGFAVSAC 420

Db 421 HVHDEFRTAAVEGPVTFLDMEDGYNIPQTDESTLMTIAYVMAAACALFMPLCLMVCQW 480

QY 421 HVHDEFRTAAVEGPVTFLDMEDGYNIPQTDESTLMTIAYVMAAACALFMPLCLMVCQW 480

Db 421 HVHDEFRTAAVEGPVTFLDMEDGYNIPQTDESTLMTIAYVMAAACALFMPLCLMVCQW 480

QY 481 RCLRLRQHDDFADDISLLK 501

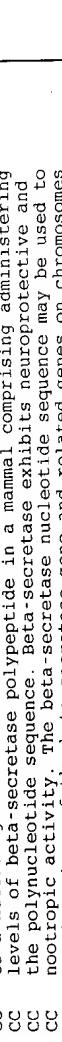
Db 481 RCLRLRQHDDFADDISLLK 501

## RESULT 10

AAB07896 standard; Protein; 501 AA.  
ID AAB07896  
XX AAB07896;  
AC AAB07896;  
XX DT 14-NOV-2000 (first entry)

XX DE Amino acid sequence of a 'human beta-secretase enzyme.  
XX DE Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;  
KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;  
KW inhibitor.  
XX DE Homo sapiens.  
OS XX WO200047618-A2.  
XX PD 17-AUG-2000.  
XX PA 10-FEB-2000; 2000WO-US03819.  
XX PR 10-FEB-1999.; 99US-0119571.  
XX PR 15-JUN-1999.; 99US-0139172.  
XX PA (ELAN) ELAN PHARM INC.  
XX PI Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;  
PI Srinivas S, Tasuno G, Tung J, Wang S, McConlogue L;  
XX DR WPI; 2000-531011/48.  
DR N-PSDB; AAA5550, AAA5550.

XX Purified beta-secretase protein used in assays to discover inhibitors  
PT which can be used for the treatment of amyloidogenic diseases e.g.  
PT Alzheimer's disease -  
XX Claim 17; Fig 2A; 121pp; English.  
XX The specification describes a beta-secretase enzyme. The enzyme cleaves  
CC beta-amyloid precursor protein to produce beta-amyloid peptide. This  
CC enzyme is therefore implicated in the production of amyloid plaque  
CC components which accumulate in the brains of individuals afflicted with  
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to  
CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's  
CC disease-like pathology to test if they maintain or improve cognitive  
CC ability or reduce the plaque burden. The compounds are used for the  
CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The  
CC present sequence represents a human beta-secretase enzyme.





DT	26-OCT-1998	(first entry)	Db	361 ILPOOYLRPEVEDATSDDCYKFAISSOSSTGTVMGAVIMEGFYVVFDARKRIGFAVSAC 420
XX	Amino acid sequence of human ASP2 (aspartic protease 2).		Qy	4.21 HVHDEFTPAVEGFVTLDMEDCGYNFQTDESTLMTIAYMAATCALFMLPLCLMVCW 480
DE			Db	4.21 HVHDEFTPAVEGFVTLDMEDCGYNFQTDESTLMTIAYMAATCALFMLPLCLMVCW 480
XX	Human; ASP2; aspartic protease 2; agonist; antagonist; immunospecific; antibody; inhibition; Alzheimer's disease; cancer; protease; prohormone processing.		Qy	4.21 HVHDEFTPAVEGFVTLDMEDCGYNFQTDESTLMTIAYMAATCALFMLPLCLMVCW 480
KW			Db	4.81 RCLRLRQHDDADDISLK 501
KW			Qy	4.81 RCLRLRQHDDADDISLK 501
KW			Db	4.81 RCLRLRQHDDADDISLK 501
XX	Homo sapiens.			
OS				
XX				
PN	EP85544-A2.			
XX	29-JUL-1998.			
PD				
XX	98EP-0300573.			
PF				
XX	27-JAN-1998;			
PR				
XX	97GB-0001684.			
PS				
XX	(SMIK ) SMITHKLINE BEECHAM CORP.			
PI	(SMIK ) SMITHKLINE BEECHAM PLC.			
XX	Chapman CG, Murphy K, Powell DJ, Smith TS;			
PX				
XX	DR; AAV809/34.			
DR				
XX	DR; AAV1696.			
PT	New nucleic acid encoding human aspartic protease 2 - used to treat			
PT	prevent and diagnose e.g. Alzheimer's disease, cancer and prohormone			
PT	processing			
XX				
PS	Claim 1; Page 7; 26pp; English.			
XX				
PS	This is the amino acid sequence of the human ASP2 (aspartic protease CC family), used in the method of the invention. Agonists and CC antagonists for ASP2 immunospecific antibodies are used to treat CC conditions requiring increased or decreased activity or expression of CC ASP2 respectively. ASP2 is used to treat and diagnose e.g. CC Alzheimer's disease, cancer and prohormone processing and ASP2 or a CC fragment can be used to induce an immune response against the above CC conditions.			
XX				
SO	Sequence 501 AA;			
Query Match	99.5%;	Score 2650; DB 19; Length 501;		
Best Local Similarity	99.6%;	Pred. No. 1.5e-261;		
Matches 499; Conservative 0; Mismatches 2; Indels 0;		Gaps 0;		
Db	1 MAQALPMLLNGAGYLPAHGTOHGIRPLRSGLGGAPLGLRPRETDEEPEPGREGSF 60		PS	Claim 20; SEQ ID NO 39970; 103pp; English.
Qy	1 MAQALPMLLNGAGYLPAHGTOHGIRPLRSGLGGAPLGLRPRETDEEPEPGREGSF 60		XX	
Db	61 VEMVDNLRGKSGQGYEMTGSPPTLNIVDGTSSNFAVGAAPHLRYTORQSLST 120		CC	The invention relates to isolated polynucleotide (I) and
Qy	61 VEMVDNLRGKSGQGYEMTGSPPTLNIVDGTSSNFAVGAAPHLRYTORQSLST 120		CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,
Db	61 VENVDNLRGKSGQGYEMTGSPPTLNIVDGTSSNFAVGAAPHLRYTORQSLST 120		CC	polymerase chain reaction (PCR) primers, oligoners, and for chromosome
Qy	121 YDLRKGVYYVPTQKGKEGLTDYSPKGPNVTVRNATAITESDKEFFITNSNWEGL 180		CC	and gene mapping, and in recombinant production of (II). The
Db	121 YDLRKGVYYVPTQKGKEGLTDYSPKGPNVTVRNATAITESDKEFFITNSNWEGL 180		CC	polynucleotides are also used in diagnostics as expressed sequence tags
Qy	121 YDLRKGVYYVPTQKGKEGLTDYSPKGPNVTVRNATAITESDKEFFITNSNWEGL 180		CC	for identifying expressed genes. (I) is useful in gene therapy techniques
Db	121 YDLRKGVYYVPTQKGKEGLTDYSPKGPNVTVRNATAITESDKEFFITNSNWEGL 180		CC	to restore normal activity of (II) or to treat disease states involving
Qy	121 YDLRKGVYYVPTQKGKEGLTDYSPKGPNVTVRNATAITESDKEFFITNSNWEGL 180		CC	(II). (II) is useful for generating antibodies against it, detecting or
Db	121 YDLRKGVYYVPTQKGKEGLTDYSPKGPNVTVRNATAITESDKEFFITNSNWEGL 180		CC	quantitating a polypeptide in tissue, as molecular weight markers and as
Qy	121 YDLRKGVYYVPTQKGKEGLTDYSPKGPNVTVRNATAITESDKEFFITNSNWEGL 180		CC	quantitating a polypeptide in tissue, as molecular weight markers and as
Db	121 YDLRKGVYYVPTQKGKEGLTDYSPKGPNVTVRNATAITESDKEFFITNSNWEGL 180		CC	a food supplement. (II) and its binding partners are useful in medical
Qy	121 YDLRKGVYYVPTQKGKEGLTDYSPKGPNVTVRNATAITESDKEFFITNSNWEGL 180		CC	CC imaging of sites expressing (II). (I) and (II) are useful for treating
Db	121 YDLRKGVYYVPTQKGKEGLTDYSPKGPNVTVRNATAITESDKEFFITNSNWEGL 180		CC	CC disorder involving aberrant protein expression or biological activity.
Qy	121 YDLRKGVYYVPTQKGKEGLTDYSPKGPNVTVRNATAITESDKEFFITNSNWEGL 180		CC	The polypeptide and polynucleotide sequences have applications in
Db	121 YDLRKGVYYVPTQKGKEGLTDYSPKGPNVTVRNATAITESDKEFFITNSNWEGL 180		CC	CC diagnostics, forensics, gene mapping, identification of mutations
Qy	121 YDLRKGVYYVPTQKGKEGLTDYSPKGPNVTVRNATAITESDKEFFITNSNWEGL 180		CC	CC responsible for genetic disorders or other traits to assess biodiversity
Db	121 YDLRKGVYYVPTQKGKEGLTDYSPKGPNVTVRNATAITESDKEFFITNSNWEGL 180		CC	CC and to produce other types of data and products dependent on DNA and
Qy	121 YDLRKGVYYVPTQKGKEGLTDYSPKGPNVTVRNATAITESDKEFFITNSNWEGL 180		CC	CC amino acid sequences. ABG0010-ABG3037 represent novel human
Db	121 YDLRKGVYYVPTQKGKEGLTDYSPKGPNVTVRNATAITESDKEFFITNSNWEGL 180		CC	CC diagnostic amino acid sequences of the invention.
Qy	121 YDLRKGVYYVPTQKGKEGLTDYSPKGPNVTVRNATAITESDKEFFITNSNWEGL 180		CC	The sequence data for this patent did not appear in the printed
Db	121 YDLRKGVYYVPTQKGKEGLTDYSPKGPNVTVRNATAITESDKEFFITNSNWEGL 180		CC	CC specification, but was obtained in electronic format directly from WIPO
Qy	121 YDLRKGVYYVPTQKGKEGLTDYSPKGPNVTVRNATAITESDKEFFITNSNWEGL 180		CC	CC at ftp.wipo.int/pub/published_pct_sequences.
Db	121 YDLRKGVYYVPTQKGKEGLTDYSPKGPNVTVRNATAITESDKEFFITNSNWEGL 180		XX	Sequence 969 AA;
Qy	121 YDLRKGVYYVPTQKGKEGLTDYSPKGPNVTVRNATAITESDKEFFITNSNWEGL 180		SO	

Query Match 97.2%; Score 2588; DB 22; Length 969;  
 Best Local Similarity 98.0%; Pred. No. 8.4e-255;  
 Matches 492; Conservative 0; Mismatches 9; Indels 1; Gaps 1;  
 XX

Qy 1 MAQALPWLLWMGAGVLPANGTQHIGIRLPLRSLGCGAPGLRLPRETDEPEEPGRGSP .60  
 Db 1 MAQALPWLLWMGAGVLPANGTQHIGIRLPLRSLGCGAPGLRLPRETDEPEEPGRGSP 60

Qy 61 VEMVDNLRGKSGQQYVEMTVGSPQTNLIVDTGSNNFAVGAAAPHFPLHRYQQLSSR 120  
 Db 61 VEMVDNLRGKSGQQYVEMTVGSPQTNLIVDTGSNNFAVGAAAPHFPLHRYQQLSSR 120

Qy 121 YRDLRKGVYVPTQYKWEQELGTDLVSIHPGPNTYVRAAATTESDKFFINGSNWEGIL 180  
 Db 121 YRDLRKGVYVPTQYKWEQELGTDLVSIHPGPNTYVRAAATTESDKFFINGSNWEGIL 180

Qy 181 GLAYAEIARPDDSLPFFDSLVKQTHVPNLFSLHLCGAGFPPLNOSEVLAvggsmiIGGI 240  
 Db 181 GLAYAEIARPDDSLPFFDSLVKQTHVPNLFSLQLCGAGFPPLNOSEVLAvggsmiIGGI. 240

241 DHSLTGSLWTPIRREWWYYEVIVRVEINGDQLKMDCKEYNYDKSIVDGGTNLRPLKK 300  
 241 DHSLTGSLWTPIRREWWYYEVIVRVEINGDQLKMDCKEYNYDKSIVDGGTNLRPLKK 300

Qy 301 VFEAAVKSIIKAASSTEKFDPGEOLQYCWQAGTTPWNIFPVISLYLGEVTNOSFRAT 360  
 Db 301 VFEAAVKSIIKAASSTEKFDPGEOLQYCWQAGTTPWNIFPVISLYLGEVTNOSFRAT 360

Qy 361 ILPOQYLRYPEVEDATSQDCYKFAISSTEKFDPGEOLQYCWQAGTTPWNIFPVISLYLGEVTNOSFRAT 420  
 Db 361 ILPOQYLRYPEVEDATSQDCYKFAISSTEKFDPGEOLQYCWQAGTTPWNIFPVISLYLGEVTNOSFRAT 420

Qy 421 HVHDEFRTAAVEGPFTVLDMDCGWNIQPQDESTLMIAVMAAICALPMLPCLMVCQN 480  
 Db 421 HVHDEFRTAAVEGPFTVLDMDCGWNIQPQDESTLMIAVMAAICALPMLPCLMVCQN 480

Qy 481 RCLRCLRQHQHD-DFADDISLKK 501  
 Db 481 RCLRCLRQHQMTLPDDISLKK 502

RESULT 14  
 ID AAB66572 standard; Protein; 488 AA.  
 XX Human memapsin 2.  
 AC AAB65572;  
 DT 12-APR-2001 (first entry)  
 XX Human memapsin 2.

KW Human; memapsin 2; nootropic; neuroprotective; amyloid precursor protein;  
 KW APP; memapsin 2 inhibitor; Alzheimer's disease.  
 OS Homo sapiens.  
 PN WO200100665-A2.  
 PD 04-JAN-2001.  
 XX 27-JUN-2000; 2000WO-US17742.  
 PR 28-JUN-1999; 99US-0141363.  
 PR 30-NOV-1999; 99US-0168060.  
 PR 25-JAN-2000; 2000US-0177836.  
 PR 27-JAN-2000; 2000US-0178368.  
 PR 08-JUN-2000; 2000US-0210492.  
 XX (OKLA-) OKLAHOMA MEDICAL RES FOUND.  
 PA (UNIT ) UNIV ILLINOIS FOUND.  
 XX Tang JIN, Hong L, Ghosh AK;  
 PI Tang JIN, Hong L, Ghosh AK;

XX DR WPI-2001-137933/14.  
 N-PSDB; AAF31848.  
 XX Novel memapsin 2 inhibitors which bind to active site of memapsin 2 having 2 catalytic aspartic residues and substrate binding cleft, used to treat Alzheimer's disease by blocking amyloid precursor protein cleavage -  
 XX Example 1; Page 72-74; 86pp; English.  
 XX The present sequence is given in a specification relating to an inhibitor of catalytically active memapsin 2. The inhibitor binds to the memapsin 2 active site, which is defined by the presence of two catalytic aspartic residues and a substrate binding cleft. The inhibitor is useful for the treatment and diagnosis of Alzheimer's disease. It is useful in screens for individuals with a genetic predisposition to Alzheimer's disease. The inhibitor is useful as a reagent for specifically binding to memapsin 2 or memapsin 2 analogues and for aiding in memapsin 2 isolation, purification and characterisation.  
 PS Sequence 488 AA;  
 XX Query Match 96.9%; Score 2582; DB 22; Length 488;  
 CC Best Local Similarity 99.8%; Pred. No. 1.3e-254;  
 CC Matches 487; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 Qy 14 AGVLDAHGTOHIGIRLPLRSLGGAFLRLPRETDEPEEPGRGGSFVENVDNLRGKSGQ 73  
 Db 1 AGVLDAHGTOHIGTRPLRSLGGAFLRLPRETDEPEEPGRGGSFVENVDNLRGKSGQ 60  
 Qy 74 GYYVETVGSPPQTNLIVDTGSNNFAVGAAAPHPLHRYQRLSTYRDLRKGVYVPTY 133  
 Db 61 GYYVETVGSPPQTNLIVDTGSNNFAVGAAAPHPLHRYQRLSTYRDLRKGVYVPTY 133  
 Qy 194 LEPPFDLSLVKQTHVNPNLFLHLCGAGFPPLQSEVAVSGGSMITGIDHSLYTGSWLWTP 253  
 Db 181 LEPPFDLSLVKQTHVNPNLFLHLCGAGFPPLQSEVAVSGGSMITGIDHSLYTGSWLWTP 253  
 Qy 254 IRRENEYEVIVRVEINGDQLKMDCKEYNYDKSIVDGGTNLRPLKKVEAAVSKAAS 313  
 Db 241 IRRENEYEVIVRVEINGDQLKMDCKEYNYDKSIVDGGTNLRPLKKVEAAVSKAAS 300  
 Qy 314 STEKPDGFELGQIYCWQAGTTPWNIFPVISLYLGEVTNOSFRITILPQQLFPLVEDY 373  
 Db 301 STEKPDGFELGQIYCWQAGTTPWNIFPVISLYLGEVTNOSFRITILPQQLFPLVEDY 360  
 Qy 374 ATSDDDCYKFAISQSSTGTVMGAVIMEGFYVVFDAKRKGFAVACHVIDEFTAAVEG 433  
 Db 361 ATSDDDCYKFAISQSSTGTVMGAVIMEGFYVVFDAKRKGFAVACHVIDEFTAAVEG 493  
 Qy 434 PFVTLDMEDGYNIQDGTNSFRITILPQQLFPLVEDY 420  
 Db 421 PFVTLDMEDGYNIQDGTNSFRITILPQQLFPLVEDY 480  
 XX RESULT 15  
 AAB61334  
 ID AAB61334 standard; Protein; 488 AA.  
 XX AC AAB61334;  
 AC AAB61334;  
 XX DE Memapsin 2 protein.  
 XX XX 02-APR-2001 (first entry)  
 XX DT 02-APR-2001  
 XX DE Memapsin 2 protein.

KW Memapsin 2; catalyst; Alzheimer's.  
 XX Homo sapiens.

XX WO200100663-A2.

XX PD 04-JAN-2001.

XX PP 27-JUN-2000; 2000WO-US17661.

XX PR 28-JUN-1999; 99US-0141363.

XX PR 30-NOV-1999; 99US-0168060.

XX PR 25-JAN-2000; 2000US-0177836.

XX PR 27-JAN-2000; 2000US-0178368.

XX PR 08-JUN-2000; 2000US-0210292.

XX PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.

XX Tang JUN, Lin X, Koelsch G;  
 WPI: 2001-102885/11.

XX Purified recombinant catalytically active memapsin 2, used to screen  
 PR inhibitors of it, which are used to treat and prevent Alzheimer's  
 PR disease -  
 XX

PS Claim 2; Page 73-75; 86pp; English.

CC The present invention relates to a purified recombinant  
 CC catalytically active memapsin 2. The invention may be used for  
 CC isolating inhibitors which are used to treat or prevent  
 CC Alzheimer's disease. The invention may also be used to screen  
 CC for individuals more genetically prone to develop Alzheimer's  
 CC disease.

XX Sequence 488 AA;

Query Match 96.9%; Score 2582; DB 22; Length 488;  
 Best Local Similarity 99.8%; Pred. No. 1, 3e-254;  
 Matches 487; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 14 AGVLPAGTOHQGIRPLRSLGLGGAPLGLRLPRETDEEPEPGRGSEVMNDNLRSKSGQ 73  
 Db 1 AGVLPAGTOHQGIRPLRSLGLGGAPLGLRLPRETDEEPEPGRGSEVMNDNLRSKSGQ 60

QY 74 GYYVEMTVGSPPQTINLYLDTGSSENFAAGAAPHPIFLRYYQRLSSYRDLRGVVPY 133  
 61 GYYVEMTVGSPPQTINLYLDTGSSENFAAGAAPHPIFLHYWYQRLSSYRDLRGVVPY 120

QY 134 QGKWEGLGFDLVSPHGPNTVYANIAAITESDPKFFNSMWEGILGLAYATEIARDDS 193  
 Db 121 QGKWEGLGFDLVSPHGPNTVYANIAAITESDPKFFNSMWEGILGLAYATEIARDDS 180

QY 194 LEPPFDLSVKOTHVYNLFSHLGAGFPLNQSEVIALSTGGSMITGGIDHSLYTGSLWTP 253  
 Db 181 LEPPFDLSVKOTHVYNLFSLQLCAGFPLNQSEVIALSTGGSMIGGIDHSLYTGSLWTP 240

QY 254 IRREWYEVIVRVEINGDQLMKCQEYNDKSTVDSGTNNLRPKYFEAAVSKTAAAS 313  
 Db 241 IRREWYEVIVRVEINGDQLMKCQEYNDKSTVDSGTNNLRPKYFEAAVSKTAAAS 300

Db 301 STEKFPPGFWLGEQIVCWAQTTPNNFPPVTSILMGETVNOSERITILPOQYLRPVNDV 373

QY 374 ATSQDDCYKFAISQSSTGTYMGAVIMEFGYVVFDRKRIGFAVSACHVDEFRTAAVEG 433

Db 361 ATSQDDCYKFAISQSSTGTYMGAVIMEFGYVVFDRKRIGFAVSACHVDEFRTAAVEG 420

QY 434 PFTVLDMEDGCGNIPQDESTLMTIAVMAICALFMPLCLMYCQWRCLRCRQHDDF 493

Db 421 PFTVLDMEDGCGNIPQDESTLMTIAVMAICALFMPLCLMYCQWRCLRCRQHDDF 480

QY 494 ADDISILK 501  
 Db 481 ADDISILK 488

Search completed: June 5, 2003, 16:12:15  
 Job time : 40 secs

34	273	10.2	398	1	PEPF_RABIT	P27823
35	270.5	10.2	381	1	CHW_CALIJ	Q9n2d2
36	268	10.1	396	1	CARP_NEUCR	Q01294
37	267	10.0	365	1	CATD_SHEEP	Q9m588
38	266.5	10.0	381	1	PEPF_CALIJ	Q9n2d3
39	266	10.0	394	1	PEPC_CAVPO	Q64411
40	266	10.0	405	1	CARP_YEAST	P07267
41	266.5	9.9	388	1	PEPF_HUMAN	P20422
42	264	9.9	388	1	PAG_HORSE	Q28389
43	262	9.8	496	1	ASPR_ORYSA	P42211
44	261.5	9.8	387	1	ASPR_AEDEA	Q03168
45	261	9.8	345	1	CATD_PIG	P00795
					sus scrofa	
ALIGNMENTS						
RESULT 1						
BACE_HUMAN		STANDARD;	PRT;	501 AA.		
ID	BACE_HUMAN					
AC	P56817; Q9UJT5; Q9BYC1; Q9BYC9;					
DT	30-MAY-2000 (Rel. 39, Created)					
DT	30-MAY-2000 (Rel. 39, Last sequence update)					
DT	15-JUN-2002 (Rel. 41, Last annotation update)					
DE	Beta-secretase precursor (EC 3.4.23.-) (Beta-site APP cleaving enzyme)					
DE	(beta-site amyloid precursor protein cleaving enzyme) (Aspartyl protease 2) (Asp 2) (Asp 2) (Membrane-associated aspartic protease 2)					
DE	(Memapsin-2), BACE OR BACE1.					
GN	OsO <sub>4</sub> (Human).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.					
OX						
RN	[1]					
RP	SEQUENCE FROM N.A. (ISOFORM A).					
RC	TISSUE-Brain;					
RC	MEDLINE=200209272; PubMed=10531052;					
RA	Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendez E.A., Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y., Fisher S., Fuller J., Burgess T., Lile J., Jarosinski M.A., Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F., Treanor J., Rogers G., Citron M.;					
RA	"Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE.";					
RA	Science 286:735-741(1999).					
RN	[2]					
RP	SEQUENCE FROM N.A. (ISOFORM A), SEQUENCE OF 46-68, AND RP CHARACTERIZATION.					
RC	MEDLINE=20051171; PubMed=10591214;					
RA	Sinha S., Anderson J.P., Barbour R., Basi G.S., Caccavello R., Denis D., Doan M., Doyle H.F., Frigon N., Hong J., Jacobson-Croak K., Jewett N., Keim P., Knops J., Lieberburg T., Power M., Tan H., Tatsuno G., Tung J., Schenk D., Seubert P., Suomensuuari S.M., Walker D., Zhao J., McConlogue L., Vargheese J.;					
RA	"Purification and cloning of amyloid precursor protein beta-secretase from human brain.";					
RA	Nature 402:537-540(1999).					
RN	[3]					
RP	SEQUENCE FROM N.A. (ISOFORM A).					
RX	MEDLINE=20051170; PubMed=10591213;					
RA	Yen R., Bierkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M., Brashier J.R., Stratman N.C., Mathews W.R., Buhl A.E., Carter D.B., Tonasselli A.G., Parodi L.A., Heinrikson R.L., Gurney M.E.;					
RA	"Membrane-anchored aspartyl protease with Alzheimer's disease beta-secretase activity".					
RL	Nature 402:537-540(1999).					
RN	[4]					
RP	SEQUENCE FROM N.A. (ISOFORM A).					
RX	MEDLINE=2012043; PubMed=10656250;					
RA	Hussain I., Powell D.J., Howlett D.R., New D.G., Meek T.D., Chapman C., Gloger I.S., Murphy K.E., Southan C.D., Carter D.B., Smith M., Cimino J., Gurney M.E., Heinrikson R.L., Gurney M.E.;					
RA	"Membrane-anchored aspartyl protease with Alzheimer's disease beta-secretase activity".					
RL	Nature 402:537-540(1999).					
RN	[5]					
RP	SEQUENCE FROM N.A. (ISOFORM A).					
RX	MEDLINE=2012043; PubMed=10656250;					
RA	Yen R., Bierkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M., Brashier J.R., Stratman N.C., Mathews W.R., Buhl A.E., Carter D.B., Tonasselli A.G., Parodi L.A., Heinrikson R.L., Gurney M.E.;					
RA	"Membrane-anchored aspartyl protease with Alzheimer's disease beta-secretase activity".					
RL	Nature 402:537-540(1999).					

RT "Identification of a novel aspartic proteinase (Asp 2) as beta-secretase";  
 RT J. Neurosci. 14:419-427(1999).  
 RL [5]  
 RP SEQUENCE FROM N.A. (ISOFORM B).  
 RC TISSUE=Brain, and Pancreas;  
 RC Michel B., De Pietri Tonelli D., Zacchetti D., Keller P.;  
 RT "New beta-site APP cleaving enzyme isoform (BACE-1B) obtained from  
 RT human brain and pancreas";  
 RL [6]  
 RP SEQUENCE FROM N.A. (ISOFORM C).  
 RC TISSUE=Pancreas; Pietri Tonelli D., Schnurbus R.;  
 RA Zacchetti D., De Pietri Tonelli D., Schnurbus R.;  
 RT "New beta-site APP cleaving enzyme isoform (BACE-1C) obtained from  
 RT human Pancreas";  
 RL [7]  
 RP SEQUENCE FROM N.A. (ISOFORMS B; C AND D).  
 RC TISSUE=Brain;  
 RX MEDLINE=21408457; PubMed=11516562;  
 RA Lin X., Koelsch G., Wu S., Downs D., Dashti A., Tang J.;  
 RT "Human aspartic protease memapsin 2 cleaves the beta-secretase site of  
 beta-amyloid precursor protein cleaving enzyme (BACE) and their effect on  
 amyloid precursor protein cleaving enzyme (BACE) and their effect on  
 amyloid beta-peptide production.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 97:1456-1460(2000).  
 RL [9]  
 RP SEQUENCE OF 14-501 FROM N.A. (ISOFORM A), AND CHARACTERIZATION.  
 RX MEDLINE=201440450; PubMed=10677483;  
 RA Fischer F., Molitari M., Bodendorf U., Paganetti P.;  
 RT "Human aspartic protease memapsin 2 cleaves the beta-secretase site of  
 beta-amyloid precursor protein.";  
 RL [9]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=21950860; PubMed=11953458;  
 RA Fischer F., Molitari M., Bodendorf U., Paganetti P.;  
 RT "The disulfide bonds in the catalytic domain of BACE are critical but  
 not essential for amyloid precursor protein processing activity.";  
 RL J. Neurochem. 80:1079-1088(2002).  
 CC -1- POSITION: RESPONSIBLE FOR THE PROTEOLYTIC PROCESSING OF THE  
 AMYLOID PRECURSOR PROTEIN (APP). CLEAVES AT THE AMINO TERMINUS OF  
 THE A-BETA PEPTIDE SEQUENCE, BETWEEN RESIDUES 671 AND 672 OF APP,  
 LEADS TO THE GENERATION AND EXTRACELLULAR RELEASE OF BETA-CLEAVED  
 SOLUBLE APP, AND A CORRESPONDING CELL-ASSOCIATED CARBOXY-TERMINAL  
 FRAGMENT WHICH IS LATER RELEASE BY GAMMA SECRETASE.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS: 4 isoforms; A/BACE-1A/BAC-501 (shown here),  
 B/BACE-1B/BACE-1A/BACE-1-457 and D/BACE-1D/BACE-1-  
 432; are produced by alternative splicing.  
 CC -1- TISSUE SPECIFICITY: BRAIN  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.ist-sib.ch/announce/>  
 CC or send an email to license@ist-sib.ch).

DR Gene7; HGNC:933; BACE.  
 DR MIM: 604252; .  
 DR InterPro: IPR001461; AspProteinaseAI.  
 DR InterPro: IPR001969; AspProteinase\_site.  
 RN Pfam: PF00026; asp\_1.  
 PRINS: PR00792; PRPSIN.  
 PROSITE: PS00141; ASP-PROTEASE\_1.  
 DR Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Transmembrane;  
 KW Signal; Alternative splicing.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT PROPEP 22 45 BETA-SECRETASE.  
 FT CHAIN 46 501 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 22 457 POTENTIAL.  
 FT TRANSMEM 458 478 CYTOPASMIC (POTENTIAL).  
 FT DOMAIN 479 501 BY SIMILARITY.  
 FT ACT\_SITE 93 93 BY SIMILARITY.  
 FT DISULFID 216 420  
 FT ACT\_SITE 289 289  
 FT DISULFID 278 443  
 FT CARBOHYD 330 380 N-LINKED (GLCNAC, . .)  
 FT CARBOHYD 153 153 (POTENTIAL).  
 FT CARBOHYD 172 172 N-LINED (GLCNAC, . .)  
 FT CARBOHYD 223 223 (POTENTIAL).  
 FT CARBOHYD 354 354 N-LINED (GLCNAC, . .)  
 FT VARSPLIC 146 189 (IN ISOFORM C AND ISOFORM D).  
 FT VARSPLIC 190 214 MISSING (IN ISOFORM B AND ISOFORM D).  
 SQ 501 AA; 55763 MW; 377FE4C824ACE05 CRC64;  
 DR Score 2656; DB 1; Length 501;  
 DR Best Local Similarity 99.8%; Pred. No. 8.3e-206;  
 DR Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 DR 1 MAQALPWLLNMGAGVLPAAHGTQHGRPLRSLGGLGAPLGRLPRETDEEPGRGSF 60  
 DR 1 MAQALPWLLNMGAGVLPAAHGTQHGRPLRSLGGLGAPLGRLPRETDEEPGRGSF 60  
 DR 61 VEMYDNLRKSGQQYYEMTVGSPQTNLVDTGSSNFAYGAAPHFLHRYYQRQLSST 120  
 DR 61 VEMYDNLRKSGQQYYEMTVGSPQTNLVDTGSSNFAYGAAPHFLHRYYQRQLSST 120  
 DR 121 YRDLRKGYYVPTQKGKWEGLDLYSPHGPNTVANIAALTEDSKFFNGSNNEGIL 180  
 DR 121 YRDLRKGYYVPTQKGKWEGLDLYSPHGPNTVANIAALTEDSKFFNGSNNEGIL 180  
 DR 181 GLAYEIAIRPDSDLEPFDSSLVKQTHVNPNSLHLGGAEPLNQSEVIALSGGSMIGGI 240  
 DR 181 GLAYEIAIRPDSDLEPFDSSLVKQTHVNPNSLHLGGAEPLNQSEVIALSGGSMIGGI 240  
 DR 241 DHSLYTGSLWYTPITREWYEVITVREVEINGDLKMDCKEYNDKIVDSGTNLRLPKK 300  
 DR 241 DHSLYTGSLWYTPITREWYEVITVREVEINGDLKMDCKEYNDKIVDSGTNLRLPKK 300  
 DR 301 VFEAAVKSIAASSEKFPQFWGEOLCQWAQTPWNPVLSLYMGEVINSFRIT 360  
 DR 301 VFEAAVKSIAASSEKFPQFWGEOLCQWAQTPWNPVLSLYMGEVINSFRIT 360  
 DR 361 ILPOQYLPRVEDVATSDQDCYKFAISQSSTGTYNGAVIMEGFYVEDARKRIGFAVSA 420  
 DR 361 ILPOQYLPRVEDVATSDQDCYKFAISQSSTGTYNGAVIMEGFYVEDARKRIGFAVSA 420  
 DR 421 HVHDERPTAAVEGPFWTLQDDGCFQVADDFDADDSLLK 501  
 DR 421 HVHDERPTAAVEGPFWTLQDDGCFQVADDFDADDSLLK 501  
 DR 481 RCLRCIQRQHDFAADDISLLK 501  
 DR 481 RCLRCIQRQHDFAADDISLLK 501  
 DR EMBL; AF201468; AAF18982; .  
 DR EMBL; AF200343; AAF17079; .  
 DR EMBL; AF204943; AAF26367; .  
 DR EMBL; AF338816; AAF338374; .  
 DR EMBL; AF338817; AAF338375; .  
 DR EMBL; AB050436; BAB04931; .  
 DR EMBL; AB050437; BAB04932; .  
 DR EMBL; AB050438; BAB04933; .  
 DR EMBL; AB050439; BAB04934; .  
 DR HSSP; P32329; 1YPS.  
 DR MEROPS; A01.0004; -.  
 AC P56819; -.  
 DR BACE\_RAT STANDARD; PRT; 501 AA.

DT	30-MAY-2000 (Rel.: 39, Created)		Db	1 MAPALRWLLWVGSSMLPAQTHLGLRPLRSGLAGPPLGLRPRETDEPEEPRRGSF 60
DT	30-MAY-2000 (Rel.: 39, Last sequence update)		QY	61 VEMDNLRGKSGQQGYVEMTGSPOTINLVDPGSNSPAGAPHFELHRYTORQLSST 120
DE	Beta-secretase precursor (EC 3.4.23.-) (Beta-site APP cleaving enzyme)		Db	61 VEMVNLRGRGSQGYVEMTGSPOTINLVDPGSNSPAGAHFELHRYTORQLSST 120
DE	(Beta-site amyloid precursor protein cleaving enzyme) (Aspartyl protease 2) (Asp 2) (Asp2) (Membrane-associated aspartic protease 2) (Memapsin-2).		QY	121 YRDLRKGVVYPYDQKWEGETDVLVSPHGPNTVTRNIAATSDKFTEFTGSNWEGIL 180
GN	BACE.		Db	121 YRDLRKGVVYPYDQKWEGETDVLVSPHGPNTVTRNIAATSDKFTEFTGSNWEGIL 180
OS	Rattus norvegicus (Rat).		QY	121 YRDLRKGVVYPYDQKWEGETDVLVSPHGPNTVTRNIAATSDKFTEFTGSNWEGIL 180
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.		Db	121 YRDLRKGVVYPYDQKWEGETDVLVSPHGPNTVTRNIAATSDKFTEFTGSNWEGIL 180
OX			QY	181 GLAYAEIARDDSLPEFFDSLVKOTHVNPFLSHLCGAGFPLNQSEVIALSGGSMIGGI 240
RN	[1]		Db	181 GLAYAEIARDDSLPEFFDSLVKOTHVNPFLSHLCGAGFPLNQSEVIALSGGSMIGGI 240
RP	SEQUENCE FROM N.A.		QY	181 GLAYAEIARDDSLPEFFDSLVKOTHVNPFLSHLCGAGFPLNQSEVIALSGGSMIGGI 240
RX	MEDLINE=20002972; PubMed=10531052;		QY	241 DHSLYTGSIMTPPIERNEYEVITVREVEINGQDLKMDCKEKNYDKSTIVDSGTTNLRLPK 300
RA	Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A., Deniz P., Teplow D.B., Ross S., Amaraite P., Loeloff R., Luo Y., Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A., Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F., Treanor J., Rogers G., Citron M.; "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE.";		Db	241 DHSLYTGSIMTPPIERNEYEVITVREVEINGQDLKMDCKEKNYDKSTIVDSGTTNLRLPK 300
CC	-!- FUNCTION: RESPONSIBLE FOR THE PROTEOLYTIC PROCESSING OF THE AMYLOID PRECURSOR PROTEIN (APP). CLEAVES AT THE AMINO TERMINUS OF THE A-BETA PEPTIDE SEQUENCE, BETWEEN RESIDUES 671 AND 672 OF APP, LEADS TO THE GENERATION AND EXTRACELLULAR RELEASE OF BETA-CLEAVED SOLUBLE APP, AND A CORRESPONDING CELL-ASSOCIATED CARBOXY-TERMINAL FRAGMENT WHICH IS LATER RELEASE BY GAMMA SECRETASE (BY SIMILARITY).		QY	301 VFEAAVKSIAASSTEKFPPDFWLGEQLYCWQAGTTPWNIFPVTSLYLGEVTNQSFRIT 360
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein.		Db	301 VFEAAVKSIAASSTEKFPPDFWLGEQLYCWQAGTTPWNIFPVTSLYLGEVTNQSFRIT 360
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY Al.		QY	361 ILPQQYLRLPVEDVATSQDDCYKFAISQSSTGTYMGAVIMEGFYYVFDARKRIGFAVSAC 420
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).		Db	361 ILPQQYLRLPVEDVATSQDDCYKFAISQSSTGTYMGAVIMEGFYYVFDARKRIGFAVSAC 420
CC	-----		QY	361 ILPQQYLRLPVEDVATSQDDCYKFAISQSSTGTYMGAVIMEGFYYVFDARKRIGFAVSAC 420
DR	EMBL; AF190727; AAC04144.1;		Db	421 HVHDFERTAAVEGPVTLMEDCGYNIPQTDESTLMTIAYVMAAICAFMLPLCLMVCON 480
DR	HSSP; P3329; 1YPS.		QY	421 HVHDFERTAAVEGPVTADMEDCGYNIPQTDESTLMTIAYVMAAICAFMLPLCLMVCON 480
DR	MEROPS; A01.004;		Db	421 HVHDFERTAAVEGPVTADMEDCGYNIPQTDESTLMTIAYVMAAICAFMLPLCLMVCON 480
DR	InterPro; IPR001461; AspProteaseAI.		QY	481 RCLRCRQHDFFADDISLIK 501
DR	InterPro; IPR001969; AspProtease.		Db	481 RCLRCRQHDFFADDISLIK 501
DR	Pfam; PF00026; asp_1.		QY	481 RCLRCRQHDFFADDISLIK 501
DR	PRINTS; PR00792; PEPSIN.		Db	481 RCLRCRQHDFFADDISLIK 501
DR	PROSITE; PS00141; ASP_Protease;		QY	481 RCLRCRQHDFFADDISLIK 501
KW	Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Transmembrane; Signal.		Db	481 RCLRCRQHDFFADDISLIK 501
FT	SIGNAL	1	QY	481 RCLRCRQHDFFADDISLIK 501
FT	PROPEP	22	21	POTENTIAL.
FT	CHAIN	46	45	POTENTIAL.
FT	DOMAIN	22	501	BETA-SECRETASE.
FT	TRANSMEM	458	457	EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	479	478	POTENTIAL.
FT	ACT_SITE	93	93	CYTOPLASMIC (POTENTIAL).
FT	ACT_SITE	289	289	BY SIMILARITY.
FT	DISULFID	216	420	BY SIMILARITY.
FT	DISULFID	278	443	BY SIMILARITY.
FT	DISULFID	330	380	BY SIMILARITY.
FT	CARBOHYD	153	153	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	172	172	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	223	223	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	354	354	N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	501 AA;	55806 MW;	24B45BC8BE87DE3 CRC64;
Query Match	96.4%		Score 2569; DB 1; Length 501;	RN [3]
Best Local Similarity	96.2%		Pred. No. 8.2e-199; Indels 0; Gaps 0;	RN [2]
Matches	482; Conservative	7; Mismatches 12;		RN [1]
Qy	1 MAQALPWLLWVGAGVLPAGTQHGLRPLRSGLGGAPLGRLPRETDEPEEPRRGGSF 60			

SEQUENCE FROM N.A.  
MEDLINE=20057170; PubMed=10591213;  
Yan R., Bienkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M., Brashier J.R., Stratman N.C., Mathews W.R., Buhl A.E., Carter D.B., Tomasselli A.G., Parodi L.A., Heiriksson R.L., Gurney M.E.; Submitted (JAN 2000) to the EMBL/GenBank/DDBJ databases.

"Membrane-anchored aspartyl protease with Alzheimer's disease beta-secretase activity";  
RT Nature 402:533-537(1999).  
RT -!  
RL FUNCTION: RESPONSIBLE FOR THE PROTEOLYTIC PROCESSING OF THE AMYLOID PRECURSOR PROTEIN (APP). CLEAVES AT THE AMINO TERMINUS OF THE A-BETA PEPTIDE SEQUENCE, BETWEEN RESIDUES 671 AND 672.  
CC LEADS TO THE GENERATION AND EXTRACELLULAR RELEASE OF BETA-SOLUBLE APP, AND A CORRESPONDING CELL-ASSOCIATED CARBOXY-  
CC FRAGMENT WHICH IS LATER RELEASE BY GAMMA-SECRETASE (BY  
CC SIMILARITY).  
CC -!  
CC SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!  
CC TISSUE SPECIFICITY: BRAIN.  
CC -!  
CC SIMILARITY: BELONGS TO PEPTIDASE FAMILY AL.  
CC -!  
CC This SWISS-PROT entry is copyright. It is produced through a  
CC between the Swiss Institute of Bioinformatics and the EMBL-  
CC European Bioinformatics Institute. There are no restrictions  
CC use by non-profit institutions as long as its content is  
CC modified and this statement is not removed. Usage by and in  
CC entities requires a license agreement (See <http://www.isb-sib.ch>).  
CC or send an email to license@isb-sib.ch).

RX MEDLINE=20144060;	PubMed=110677483;		Matches 239; Conservative 82; Mismatches 165; Indels 33; Gaps 9;
RA Lin X., Koelisch G., Wu S., Downs D., Dashti A., Tang J.;	"Human aspartic protease memapsin 2 cleaves the beta-secretase site of beta-amyloid precursor protein";	Qy 2 AQAIPWILLWW --GAGVLPAAHCTQHGIRLPLRSGLG-----GAPL----GLR 42	
RT PROC. NATL. ACAD. SCI. U.S.A. 97:1456-1460(2000).	[7]	Db 7 ALLPLLIQLWLRAPELAPAPIT----LPLRVAAATNRVVAAPTPGDTPLRERHAOGA 61	
RN SEQUENCE FROM N.A.			
RX MEDLINE=20289799;	PubMed=10830953;	Qy 43 LPRTDEEPEPGRGSFVEMVDNLRGKSGOGYYVEMTVGSPQPTLNLYDVGSSNFAVG 102	
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,		Db 62 LALE--PALASPAQANFLAMYLNQDGSRGYRLEMIGTQPQKLQYDVGSSNFAVA 119	
RA Park H.-S., Toyoda A., Ishii K., Torikai Y., Taudien S., Blechschmidt K., Soeda E.,		Qy 103 AAPHPLFLRYQQRQSLSTYRDLRKGVVYPTQCKWEGBLGTDLVSIHPGPNTVRANIA 162	
RA Okni M., Takagi T., Sakaki Y., Lehnmann R., Patterson D., Polley A.,		Db 120 GTPHSYIDTYFDTFRSSSYRSKGDFDTVKYTQGSTWTGVGEDELVTIKGENTSEFLVNFIAT 179	
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D., Polley A.,		Qy 163 ITBESKFEFTNGSWEGIUGLAYTAIARDSDSLEFFEDSVLKOTHVPNIFSLHICAGGPL 222	
RA Reichwald A., Rump K., Schilbach M., Schudy A., Zimmermann W.,		Db 180 IFESENENFFLPFGIKWNGIUGLAYATLAKPSSSLETFDSLVTQANIPNVFSMOMGAGLPLV 239	
RA Rosenthal A., Kucioh J., Shibusawa K., Kawasaka K., Asakawa S.,		Qy 223 NQSEFLASVGGSMITGGIDHSLYTGSLWYTPIRREMYEVILVREYVQIEILKLEIGQSINLDCREVN 282	
RA Shintani A., Sasaki T., Nagamine K., Mitsuyma S., Antonarakis S.E.,		Db 240 AGS--GTNGGSLVGLGPLEPSLYKGDIWTPKEWYQIEILKLEIGQSINLDCREVN 296	
RA Minoshima S., Shimizu N., Nordsiek L., Hornischer K., Brandt P.,		Qy 283 YDKSIVDSGTTNLRPKVYERAKWSKAASSTEKFDPGFWLGEQLCYWQAGTTPNWLF 342	
RA Schärfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bioecker H.,		Db 297 ADKAVDSGTTLRLPQRFVDAVEAVARASLPIEFSDGFWGTSQALCWTNSSETPWSYFP 356	
RA Ramser J., Beck A., Klages S., Hennig S., Riesseleman L., Dagand E.,		Qy 343 VISLYLMGEVTSNFSRITLPLQQPLRPVEDVATSQDDCYKFAISQSSTGTVNGAVIMBGF 402	
RA Wehrmeyer S., Bozycz K., Gardiner K., Nizetic D., Francis F.,		Db 357 KISTYLRDENSSSRFRITLPLQYTOPMGAGLN--ECYRFGISPSSTNALVIGTVMEGF 415	
RA Lehrach H., Reinhardt R., Yaspo M.-L.,	"The DNA sequence of human chromosome 21."	Qy 403 YYVEDRARKRIGFASACHVHDERFTAAVEGPFVTLMDCGYNIPQDESTLMTIAYM 462	
RN 405:311-319(2000).		Db 416 YYFIDRAQKRQVGAASPAETAGAVSETSGPSTEDASNCVPAQSLSLEPLWIVSTAL 475	
RN SEQUENCE FROM N.A.			
RP			
RC			
RA			
RA			
RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.			
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.			
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC the European Bioinformatics Institute. There are no restrictions on its			
CC use by non profit institutions as long as its content is in no way			
CC modified and this statement is not removed. Usage by and for commercial			
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/			
CC or send an email to license@isb-sib.ch).			
CC			
DR EMBL; AF200342; AAF17078.1; -.		RESULT 5	
DR EMBL; AF117832; ADD5240.1; -.		PEP1_GADMO	
DR EMBL; AF050171; ADD45963.1; -.		ID PEP1_GADMO	
DR EMBL; AF178532; AAF29494.1; -.		DT P52472;	
DR EMBL; AF204944; AAF6368.1; -.		DT 15-JUL-1998 (Rel. 36, Created)	
DR EMBL; AF200192; AAF13714.1; -.		DT 15-JUL-1998 (Rel. 36, Last sequence update)	
DR EMBL; AL163384; CAB90458.1; -.		DT 15-JUN-2002 (Rel. 41, Last annotation update)	
DR EMBL; BC013285; CAB90554.1; -.		DE DE	
DR EMBL; BC014453; AAH14453.1; -.		OS OS	
DR HSSP; P00797; 2REN.		OC OC	
DR Genew; HGNC:934; BACE2.		OC Actinopterygia; Neopterygii; Teleostei; Euteleostomi; Neoteleoste;	
DR MIM: 605668; -.		OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.	
DR InterPro; IPR01161; Asprotease1.		OC NCBI_TaxID=8049;	
DR InterPro; IPR001969; Asprotease1.		OX RN [1]	
DR Pfam; PF00026; asp; 1.		RP SEQUENCE, AND X-RAY CRYSTALLOGRAPHY.	
DR PRINTS; PR00792; PEPSIN.		RC TISSUE=Stomach; Olsen R.L.	
DR PROSITE; PS00141; ASP_Protease; 2.		RA Karslen S., Hough E., Olsen R.L.	"Structure and proposed amino-acid sequence of a pepsin from Atlantic
KW Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Transmembrane; Signal.		RT cod (Gadus morhua)."; RT	cod (Gadus morhua)."; RT
KW SIGNAL 1 20 POTENTIAL.		RL Acta Crystallogr. D 54:32-46(1998).	Acta Crystallogr. D 54:32-46(1998).
FT PROPEP 21 ? POTENTIAL.		CC PRNTS; PR00792; PEPSIN.	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
FT DOMAIN 21 518 BETA SECRETASE 2.		DR PROSITE; PS00141; ASP_Protease; 2.	PDB; IAMS; 24-DEC-97.
FT TRANSMEM 474 473 EXTRACELLULAR (POTENTIAL).		DR Hydrolase; Aspartyl protease; Digestion; 3D-structure.	InterPro; IPR001461; Asprotease1.
FT DOMAIN 495 518 CYTOPLASMIC (POTENTIAL).		FT ACT SITE 32 32 BY SIMILARITY.	InterPro; IPR001969; Asprotease_site.
FT ACT_SITE 110 110 BY SIMILARITY.		FT ACT SITE 214 214 BY SIMILARITY.	DR PFM; PF00026; asp; 1.
FT CHAIN 303 303 BY SIMILARITY.		FT DISULFID 45 50 BY SIMILARITY.	DR PRNTS; PR00792; PEPSIN.
FT CARBOHYD 170 170 N-LINKED (GLCNAC. . . ) (POTENTIAL).		FT DISULFID 206 209 BY SIMILARITY.	DR PROSITE; PS00141; ASP_Protease; 2.
FT CONFLICT 36 36 N-LINKED (GLCNAC. . . ) (POTENTIAL).		FT DISULFID 247 280 BY SIMILARITY.	KW Hydrolase; Aspartyl protease; 3D-structure.
SQ SEQUENCE 518 AA; 56180 MW; 2E903150823760D3 CRC64;		SQ SEQUENCE 324 AA; 34014 MW; BE3A6097B694.1DD7 CRC64;	Best Local Similarity 44.18; Score 1173.5; DB: 1; Length 518; Pred. No. 1.2e-86; 1.

Page 6

INVOLVING PHENYLALANINE AND LEUCINE ARE PREFERRED, MANY OTHERS ARE ALSO CLEAVED TO SOME EXTENT.

-!- CATALYTIC ACTIVITY: Preferential cleavage: hydrophobic, preferably aromatic, residues in P1 and P1' positions. Cleaves 1-Phe-1'-Val-2, 4-Gln-1-His-5, 13-Glu-1-Ala-14, 14-Ala-1-Tyr-15, 15-Leu-1-Phe-15 and 25-Phe-1-Tyr-26 bonds in the B chain of insulin.

-!- DEVELOPMENTAL STAGE: PEPTINOGENS IN GROUP I, II, AND III WHERE THE PREDOMINANT ZYMOGENS AT LATE POSTNATAL STAGE.

-!- MISCELLANEOUS: THE EXPRESSION OF PEPTINOGEN GENES IS REGULATED BY HORMONES AND RELATED SUBSTANCES.

-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.

DR P00791; B38302.

DR MEROPS; A01.001; -.

DR InterPro; IPR001461; AsproteaseA1.

DR Pfam; PF00026; asp; 1.

PRINTS; PRO0752; PPSIN.

PROSITE; PS00141; ASP-PROTEASE; 2.

Hydrolase; Aspartyl\_protease; Digestion; zymogen; Signal; Phosphorylation; Multigene family.

KW SIGNAL: 15

FT ACTIVATION PEPTIDE: 59

FT PROPPP: 16

FT CHAIN: 60

FT MOD\_RES: 129

FT ACT\_SITE: 93

FT ACT\_SITE: 276

FT DISULFID: 106

FT DISULFID: 267

FT DISULFID: 310

FT SIGNAL: 15

FT PEPSIN: 387

FT PHOSPHORYLATION (POTENTIAL): 129

FT BY SIMILARITY: 93

FT BY SIMILARITY: 111

FT BY SIMILARITY: 111

FT BY SIMILARITY: 271

FT BY SIMILARITY: 343

FT BY SIMILARITY: 387

SQ SEQUENCE: A6EC48F7/1554.1AA48 CRC64;

Query Match 11.6%; Score 309; DB 1; Length 387; Best Local Similarity 27.1%; Pred. No. 2.3e-17; Matches 98; Conservative 66; Mismatches 130; Indels 66; Gaps 15;

Db 75 YYVETVGGPPQTNLNLVDTGSSNFAVG---AAPHPFLHRYYQRLSSSTYRDLRKGVYV 130

Db 75 YFGTISIGTPQEFTVIFDTGSSNLNWVPTSYCSSIACFLHKRFNPDDSTFOATSETLSI 134

Qy 131 PYTQGKWEGLGTDLVSIHGPNTVRAINTAESD--KFENGNSNVEGILGLAYAEI 187

Db 135 TYGTGSMTGLGYDVKV--GNIEDTNQFLGSKTEPGLTFLV-APPDGILGLAYPSI 189

Qy 188 ARPDSDLEPPSLVYKQTHY-PLNFSLHCAGGFLPQNESEVLASVGSMTIGGIDHSLYT 246

Db 190 SASDAT-PPFDNNWNNEGLYSEDLSVYLSNG-----EKGSMMFMGGDSSYT 237

Qy 247 GSLWTPIRREWWYEVILVREVEINGQDLKM-DCKEYNNYDKSIVDSGTNLRPKKKVEEA 304

Db 238 GSLNWPVPSHEGYWLTMDSTSITGETIACDSC----QAVVTGTTSLLAGPSAIK 291

Qy 305 AVKSIKAASSTEKPDGFWLGEQLV-CWQAGTTPWNIFPVISLYLMGENTNQSFRITLP 363

Db 292 IQSYLGASKNL-----LGENIISCAISDLSPDIV-----TINN 325

Qy 364 QYFLREVED-VATSOBDC--VMGAVIMEGTVVVDFARRERIGEFAV 417

Db 326 VQYPLPASAYILKEKDDCLSGFDGMNLDTSYGELMILGDFVIROQFTVFDANNQVGLAA 385

Qy 418 SA 419

Db 386 AA 387

Pepsin A-4 precursor (EC 3.4.23.1) (Pepsin I/II).

DE Pepsin A-4 precursor (EC 3.4.23.1) (Pepsin I/II).

GN Macca fuscata fuscata (Japanese macaque).

OS Macca fuscata fuscata (Japanese macaque).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Ceropithecoidea; Macaca.

OC NCBiL TaxID: 9543;

OX [1]

RN SEQUENCE FROM N.A., AND SEQUENCE OF 16-70.

RP TISSUE=Gastric mucosa;

RC MEDLINE=92037645; PubMed=1935977;

RX Kagiyama T.; Tanabe K.; Koiwai O.;

RA "Development-dependent expression of isozymogens of monkey pepsinogens and structural differences between them." ; Eur. J. Biochem. 202: 205-215(1991).

RT FUNCTION: Shows particularly broad specificity; although bonds involving phenylalanine and leucine are preferred, many others are also cleaved to some extent.

RL CATALYTIC ACTIVITY: Preferential cleavage: hydrophobic, preferably aromatic, residues in P1 and P1' positions. Cleaves 1-Phe-1-Val-2, 4-Gln-1-His-5, 13-Glu-1-Ala-14, 14-Ala-1-Tyr-15, 15-Leu-1-Tyr-16, 16-Tyr-1-Leu-17, 23-Gly-1-Phe-24, 24-Phe-1-Phe-25 and 25-Phe-1-Tyr-26 bonds in the B chain of insulin.

CC CATALYTIC ACTIVITY: Preferential cleavage: hydrophobic, preferably aromatic, residues in P1 and P1' positions. Cleaves 1-Phe-1-Val-2, 4-Gln-1-His-5, 13-Glu-1-Ala-14, 14-Ala-1-Leu-15, 15-Leu-1-Tyr-16, 16-Tyr-1-Leu-17, 23-Gly-1-Phe-24, 24-Phe-1-Phe-25 and 25-Phe-1-Tyr-26 bonds in the B chain of insulin.

CC MISCELLANEOUS: THE EXPRESSION OF PEPTINOGEN GENES IS REGULATED BY HORMONES AND RELATED SUBSTANCES.

CC PEPSIN AT PH 2.0 IN PART AS A RESULT OF THE RELEASE OF A 47 AA ACTIVATION SEGMENT AND IN PART AS A RESULT OF STEWWISE PROTEOLYTIC CLEAVAGE VIA AN INTERMEDIATE FORM(S).

CC SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: X59753; CAA42425.1; -.

DR PIR: S16065; S16065.

DR PIR: S19682; S19682.

DR HSSP: P00790; 1PSN.

DR MEROPS: A01.001; -.

DR InterPro: IPR001461; AsproteaseA1.

DR InterPro: IPR001969; Asprotease\_site.

DR InterPro: IPR001791; Laminin\_G.

DR Pfam: PF00026; asp; 1.

DR PRINTS: PRO0072; PPSIN.

DR PROSITE: PS00141; ASP-PROTEASE; 2.

DR Hydrolase; Aspartyl\_protease; Digestion; Zymogen; Multigene family; KW Signal; Glycoprotein.

FT SIGNAL: 1

FT PROPEP: 16

FT PROTOP: 39

FT CHAIN: 63

FT ACT\_SITE: 94

FT ACT\_SITE: 277

FT DISULFID: 107

FT DISULFID: 268

FT DISULFID: 311

FT CARBOHYD: 88

FT SEQUENCE: 388 AA; 41955 MW; A2223ABLF7FCDB9 CRC64;

Query Match 11.5%; Score 307.5; DB 1; Length 388;

Best Local Similarity 27.6%; Pred. No. 3.1e-17; Matches 108; Conservative 65; Mismatches 135; Indels 83; Gaps 17;

Qy 44 PRETDEEPEPGRRSSFVERVNLDLKGSGGGYVMTGSPPQTINILVDTGSSNFAVGA 103

Db 60 PTLIDEQPLE-----NYLDV-----EYFGTIGTGPQANFTVVFDTGSSNLWV - 102

Qy 104 APHPFL-----HRYYQRQLSSTYDLRKGVYVPTYQKWEGEI GTDLVSIPHGPNTV 156



FT SIGNAL	1	16	POTENTIAL.
FT CHAIN	17	383	EMBRYONIC PEPSINogen.
FT ACT_SITE	94	94	BY SIMILARITY.
FT ACT_SITE	276	276	BY SIMILARITY.
FT DISULFID	107	112	BY SIMILARITY.
FT DISULFID	267	271	BY SIMILARITY.
FT DISULFID	310	344	BY SIMILARITY.
FT CARBOHYD	132	132	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	204	204	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	309	309	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	350	350	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT	51	51	T -> S.
SQ SEQUENCE	383 AA:	41719 MW:	1642796871611F54 CRC64;
Query Match		11.38;	Score 301.5; DB 1; Length 383;
Best Local Similarity	25.28;	Pred. No. 9.1e-17;	
Matches 90;	Conservative 76;	Mismatches 124;	Indels 67; Gaps 14;
Qy	75 YYVEMTVGSPQQTQLNLVDFGSSNFAVGA---APHPFLHRYYQFOLSTYRDLRKGVYVW 130		
Db	193 -DGTPVFDNMVNNEQNLFLSVL-----REPMGSWVFGCGIDESYFTGSI 240		
Qy	76 YYGTISIGTPQDFDTIVPFDGSSNLWVPVSCTSPACQSHQMNFNSQSTYKSTGQNLSI 135		
Qy	131 PYTOGKWEGLGTDLVSPHGPNTVTRANIAAITESDKFIRRINGSNWEGTGLAYAEIARP 190		
Db	136 HYGTGDMEGTIVGCDIVTVAISLMDTNQFLGJST-SBPGQFFVY-YKFDGTLGLCPSLAA- 192		
Qy	191 DDSLEPPEDSLVKQTHV-PNLFSLHLCGAGEPLNOSEVLAvgGSMIIGGIDHSLYTGS 249		
Db	193 -DGTPVFDNMVNNEQNLFLSVL-----REPMGSWVFGCGIDESYFTGSI 240		
Qy	250 WYTPIRREWVYEVILVRIEVEQGDL---KMPKEYNDKSIIVDSGTTNLRLPKKVFEAAV 307		
Db	241 NWIPISYQGYWQISMSDIINNKQEACASSGC----QALIDGTSVLAGPASDINDQS 294		
Qy	308 SIIKAASSTKEFPDGFWLGEOLVCNQAGTPTPNWIFPVISL---YLMGEVTNOSFRITILP 363		
Db	295 AVGADONT-----YEYGSV-----NCSHILAMPMDVVPTIGG----- 326		
Qy	364 QQYLRPVEDVVA---TSQDCYKFAISQSSGTGVNGAVIMEGFYVVFDDARKRKGFA 416		
Db	327 -QY--PVALAYTEQNGQTCMSSFQNSADDWLWLGDFVIRYVSSIFDRANNRVLGA 380		
RESULT 11			
ID CATE_HUMAN	STANDARD;	PRT;	396 AA.
AC P14091;			
Y 01-JAN-1990	(Rel. 13, Created)		
D 01-JAN-1990	(Rel. 13, Last sequence update)		
D 15-JUN-2002	(Rel. 41, Last annotation update)		
DE Cathepsin E precursor (EC 3.4.23.34).			
GN CTSE.			
OS Homo sapiens (Human).			
OC Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates;			
OX NCBI_TAXID=9606;			
RN [1] Sequence from N.A.			
RX MEDLINE=89380302; PubMed=267141;			
RX MEDLINE=9212877; PubMed=1370478;			
RA Azuma T., Liu W.G., Mohandas T.K., Couvreur J.M., Taggart R.T.;			
RT "Human gastric cathepsin E gene. Multiple transcripts result from alternative polyadenylation of the primary transcripts of a single gene locus at 1q31-q32.";			
RT J. Biol. Chem. 267:16748-16753(1992).			
RL [2] J. Biol. Chem. 264:16748-16753(1999).			
RP SEQUENCE FROM N.A.			
RP MEDLINE=9212877; PubMed=1370478;			
RA Azuma T., Liu W.G., Vander Laan D.J., Bowcock A.M., Taggart R.T.;			
RT "Human gastric cathepsin E gene. Multiple transcripts result from alternative polyadenylation of the primary transcripts of a single gene locus at 1q31-q32.";			
RT J. Biol. Chem. 267:1609-1614(1992).			
RN [3] J. Biol. Chem. 264:16748-16753(1999).			

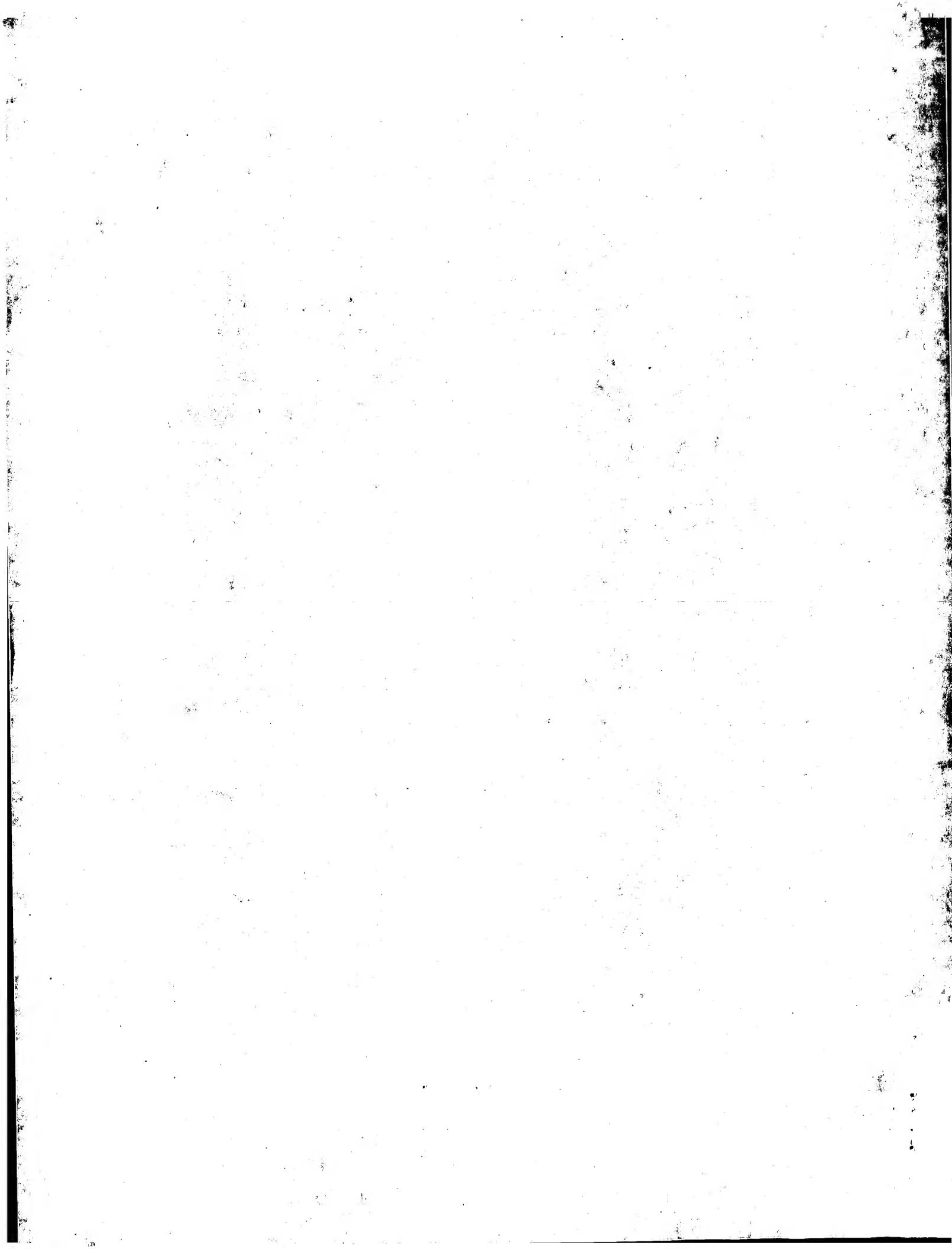
RESULT 12		S19aa and active site.
CATID_HUMAN	STANDARD;	PRT; 412 AA.
P07339; AC		
DT 01-APR-1988 (Rel. 07, Created)		
DT 01-APR-1988 (Rel. 07, Last sequence update)		
DT 15-JUN-2002 (Rel. 41, Last annotation update)		
CATHESIN D precursor (EC 3.4.23.5).		
CNSD.		
OS 'Homo sapiens' (Human).		
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC OCNCBI_TaxID=9606;		
OX NCBI_TaxID=9606;		
RN SEQUENCE FROM N.A.		
RX MEDLINE=85270436; PubMed=3927729;		
RX Faust P.L., Kornfeld S., Chirgwin J.M.;		
RX "Cloning and sequence analysis of cDNA for human cathepsin D.";		
RX Proc. Natl. Acad. Sci. U.S.A. 82:4910-4914 (1985).		
[2]		
RN SEQUENCE FROM N.A.		
RX MEDLINE=97231068; PubMed=35988310;		
RX Westley B.R., May F.E.B.;		
RX Redecker B., Heckendorf B., Grossch H.W., Mersmann G., Hasilik A.;		
RX "Oestrogen regulates cathepsin D mRNA levels in oestrogen responsive		
human breast cancer cells";		
Nucleic Acids Res. 15:3773-3786(1987).		
[3]		
RN SEQUENCE FROM N.A.		
RX MEDLINE=911299158; PubMed=2069717;		
RX Redecker B., Heckendorf B., Grossch H.W., Mersmann G., Hasilik A.;		
RX "Molecular organization of the human cathepsin D gene.";		
RX DNA Cell Biol. 10:423-431(1991).		
[4]		
RN SEQUENCE FROM N.A.		
RX TISSUE=Kidney;		
RX Strausberg R.;		
RX Submitted (Oct'2001) to the EMBL/GenBank/DBJ databases.		
RN [5]		
RX SEQUENCE OF 1-222 FROM N.A.		
RX MEDLINE=94085791; PubMed=8262386;		
RX May F.E., Smith D.J., Westley B.R.;		
RX "The human cathepsin D-encoding gene is transcribed from an estrogen-regulated and a constitutive start point.";		
RX Gene 134:277-282 (1993).		
RN [6]		
RX SEQUENCE OF 1-222 FROM N.A.		
RX MEDLINE=95021301; PubMed=7935485;		
RX Augereau P., Miralles F., Cavailles V., Gaudelet C., Parker M., Rochelefort H.		
RX RA		
RN [10]		
RN RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).		
RP TISSUE=Liver.		
RC MEDLINE=93342076; PubMed=8393577;		
RA Baldwin E.T., Bhat T.N., Gulnik S., Hosur M.V., Sowder R.C. II,		
RA Cachau R.E., Collins J., Silva A.M., Erickson J.W.;		
RT "Crystal structures of native and inhibited forms of human cathepsin D: implications for lysosomal targeting and drug design.";		
RT Proc. Natl. Acad. Sci. U.S.A. 90:6796-6800(1993).		
RL -!- FUNCTION Acid protease active in intracellular protein breakdown.		
CC Involved in the pathogenesis of several diseases such as breast		
CC cancer and possibly Alzheimer's disease.		
CC -!- CATALYTIC ACTIVITY: Specificity similar to, but narrower than,		
CC that of Pepsin A. Does not cleave the 4-Gln-1-His-5 bond in B		
CC chain of insulin.		
CC -!- SUBUNIT: CONSISTS OF A LIGHT CHAIN AND A HEAVY CHAIN.		
CC -!- SUBCELLULAR LOCATION: Lysosomal.		
CC -!- POLYMORPHISM: The Val-58 allele is significantly overrepresented		
CC in demented patients (11.8%) compared with nondemented controls (4.9%). Carriers of the Val-58 allele have a 3.1-fold increased		
CC risk for developing AD than noncarriers.		
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY AL.		
CC This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC between the Swiss Institute of Bioinformatics and the EMBL outstation		
CC at the European Bioinformatics Institute. There are no restrictions on its		
CC use by non-profit institutions as long as its content is in no way		
CC modified and this statement is not removed. Usage by and for commercial		
CC entities requires a license agreement (See http://www.isb-sib.ch/announces		
CC or send an email to license@isb-sib.ch).		
CC EMBL; M11233; AAB59529.1; -.		
DR DR EMBL; X05344; CAA28355.1; -.		
DR DR EMBL; M63138; AAA51922.1; JOINED.		
DR DR EMBL; M63134; AAA51922.1; JOINED.		
DR DR EMBL; M63135; AAA51922.1; JOINED.		
DR DR EMBL; M63136; AAA51922.1; JOINED.		
DR DR EMBL; M63137; AAA51922.1; JOINED.		
DR DR EMBL; BC016320; AAH16320.1; -.		
DR DR EMBL; L12980; AAA16320.1; -.		
DR DR EMBL; S74689; AAD14156.1; -.		
DR DR EMBL; S55357; RAD13868.1; -.		
DR DR EMBL; A25771; KHHUD.		
DR DR PUB; ILYA; 31-JAN-94.		
DR DR PUB; ILYB; 31-JAN-94.		
DR DR MEROPS; A01.009; -.		
DR DR SWISS-2DPAGE; P07339; HUMAN.		

DR	Siena-2DPAGE; P07339;	FT	TURN	343	344
Gene	HGNC: 2539; CTSD.	FT	STRAND	345	349
MIM	116840; -	FT	HELIX	351	354
DR	InterPro: IPR001461; AsproteaseAI.	FT	STRAND	355	357
DR	InterPro: IPR001969; Asprotease_site.	FT	TURN	359	362
PFam	PF00026; asp_1.	FT	STRAND	365	368
PRINTS	PR00732; PPESSIN.	FT	STRAND	370	372
PROSITE	PS00141; ASP_PROTEASE; 2.				
DR	Hydrolase; Aspartyl_protease; Glycoprotein; Lysosome; Signal; zymogen;				
KW	Polymorphism; Alzheimer's disease; 3D-structure.				
FT	SIGNAL 1 18				
PROPEP	ACTIVATION PEPTIDE.				
FT	PROPEP 19 64				
FT	CHAIN 65 412				
FT	CATHPSIN D.				
FT	CHAIN 65 161				
FT	LIGHT CHAIN (PROBABLE).				
FT	CHAIN 169 412				
FT	HEAVY CHAIN (PROBABLE).				
FT	ACT_SITE 97 97				
FT	ACT_SITE 295 295				
FT	DISULFID 91 160				
FT	DISULFID 110 210				
FT	DISULFID 286 290				
FT	DISULFID 329 366				
FT	CARBONYD 134 134				
FT	CARBONYD 263 263				
FT	VARIANT 58 58				
FT	A -> V (ASSOCIATED WITH INCREASED RISK IN AD; POSSIBLY INFLUENCES SECRETION AND INTRACELLULAR MATURATION; IN DBSNP:1751).				
FT	/FTId=VAR_011621.				
FT	STRAND 67 74				
FT	TURN 75 77				
FT	STRAND 78 85				
FT	TURN 86 89				
FT	STRAND 90 97				
FT	TURN 98 99				
FT	STRAND 103 107				
FT	TURN 108 109				
FT	HELIX 112 113				
FT	TURN 115 118				
FT	TURN 119 119				
FT	STRAND 123 123				
FT	HELIX 125 127				
FT	TURN 129 130				
FT	STRAND 132 132				
FT	STRAND 146 158				
FT	STRAND 172 184				
FT	HELIX 188 192				
FT	STRAND 197 200				
FT	HELIX 204 206				
FT	HELIX 214 220				
FT	TURN 221 222				
FT	TURN 228 233				
FT	STRAND 243 247				
FT	TURN 248 248				
FT	HELIX 254 254				
FT	STRAND 255 263				
FT	STRAND 267 267				
FT	TURN 268 269				
FT	STRAND 270 279				
FT	TURN 280 281				
FT	STRAND 284 285				
FT	TURN 287 288				
FT	STRAND 290 294				
FT	TURN 296 297				
FT	STRAND 298 298				
FT	STRAND 301 303				
FT	HELIX 305 315				
FT	TURN 316 316				
FT	STRAND 318 319				
FT	STRAND 322 323				
FT	STRAND 325 328				
FT	HELIX 329 334				
FT	STRAND 338 342				
		Query Match	11.3%	Score 300.5;	DB 1; Length 412;
		Best Local Similarity	26.9%;	Pred No. 1.2e-16;	
		Matches	123;	Mismatches	97; Gaps 21;
		Matches	123;	Conservative	68; Mismatches
		Db	5 LPWLJWWMGAGVLPAGHTONGIRPLR-----SLGGAGAPLGL-----RLP 44	Qy	
		Db	7 LPLAUCLLAA---PASAL---VRIPHAKFTSTRPMEVGSEVGSVEDLIAKGPVSKYSQAVP 60	Qy	
		Db	45 RETDDEEPEEPQRGGSFVEMDNLRKSGQQYYVEMTGVSPQTLNLLVDTGSSNFAVGAA 104	Qy	
		Db	61 AVTE-----GPIPEVKNYMDAQ---YYGEIGCTPQCFCVVFDIGSSNLWVPSI 108	Qy	
		Db	105 PHPFL-----HRYVQLQLSSTYRDLRKGVVYVTPQKGWSGELGTDVLSP-----149	Qy	
		Db	109 HCKLLDIACTWHKYNDSKSSTYVNGTSPDIHYSGSLSGYSQDTSVPCQASSASA 168	Qy	
		Db	150 HGPNTVTRVTAATTEDSKFFINGSNWEGLTGLAYAEIARPDDSFSLVKOTHV-P 208	Qy	
		Db	169 LGGVVVERQVEGAFTKQPGBTIAFKDGLIGMAVPRIS--VNNNLPVFDNLMQKLVDQ 226	Qy	
		Db	209 NLFSLHLICGAGFPNLOSEVLAvgGSMITGGIDHSLYTGSIWYTPITRREMYEVILVRE 268	Qy	
		Db	227 NIFSYL-----SRPDPAQPGGELMLGCTDSKYKGSLSYLNVTRKAYWQVHLDQE 278	Qy	
		Db	269 I-NGQDLKMDCKEYNYDKS1IVDSGTTNLRPKKVPAAVKSIAASSTEKEPDGFNLGEQ 327	Qy	
		Db	279 VASGTL-----CKE-GCEAVDTGTSLMVGPDVDEVRELQRAIGAYPLIQ-----325	Qy	
		Db	328 LV-CWQAGTTPWNIFPVVISLYLMGEVNTQSFRITILPQOYLRPVEDVATSDQCYKFAIS 386	Qy	
		Db	326 MIPCEKVST----LEAITLKLGG---KGKFLS---PEDTLVKSOAGKTL---CLSGFMG 372	Qy	
		Db	387 Q-----SSTGTVMGAVIMEGFYVVFDRARKRIGEVSA 419	Qy	
		Db	373 MDIPPPSGPLWILGDFIGRYTFEDRDNNEVGFEEAA 410	Qy	
		RESULT 1.3			
		PEP2_RABBIT			
		ID_PEP2_RABBIT			
		SEQUENCE FROM N.A.			
		RN			
		RP			
		RX			
		AC	P27B21;		
		DT	01-AUG-1992 (Rel. 23, Created)		
		DT	01-AUG-1992 (Rel. 23, Last sequence update)		
		DT	15-JUN-2002 (Rel. 41, Last annotation update)		
		DE	Pepsin II-2/3 precursor (EC 3.4.23.1) (Pepsin A).		
		OS	Oryctolagus cuniculus (Rabbit).		
		OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagidae; Oryctolagidae; Leporidae; Lagomorpha; Leporidae; Oryctolagidae.		
		NCBI_TaxID	9986;		
		RN	SEQUENCE FROM N.A.		
		SEQUENCE FROM N.A.	MEDLINE=91009127; PubMed=2129536;		
		RN	Kageyama T., Tanabe K., Koawai O.		
		RT	"Structure and development of rabbit pepsinogens. Stage-specific zymogens, nucleotide sequences of cDNAs, molecular evolution, and gene expression during development."		
		RT	J. Biol. Chem. 265:17031-17038 (1990).		
		CC	-!- FUNCTION: SHOWS PARTICULARLY BROAD SPECIFICITY; ALTHOUGH BONDS INVOLVING PHENYLALANINE AND LEUCINE ARE PREFERRED, MANY OTHERS ARE ALSO CLEAVED TO SOME EXTENT.		
		CC	-!- CATALYTIC ACTIVITY: Preferential cleavage: hydrophobic, preferably aromatic, residues in P1 and P1'. Cleaves 1-Phe-1-val-2, 4-Gln-1-His-5, 13-Glu-1-Ala-14, 14-Ala-1-Leu-15, 15-Leu-1-Tyr-16, 16-Tyr-1-Leu-17, 23-Gly-1-Phe-24, 24-Phe-1-Phe-25 and 25-Phe-1-		
		CC	CC bonds in the B chain of insulin.		



QY	354	NQSFRITILPQQQLRPPVEDVATSSDDCYKFAISOSSTGT-----VMGAVIMEGFYVVYFD	407		KW	Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Signal;
Db	322	-----TINGIQYPLPASAYILKEKEDDCTSGLEGMNVDITYGELWILDGVFROYFTVFD	375		KW	POTENTIAL.
AC	P24268				FT SIGNAL	1
AC	CATD_RAT	STANDARD;	PRT:	4.07 AA.	FT PROPEP	21
DT	01-MAR-1992	(Rel. 21, Created)			FT CHAIN	65
DT	15-JUN-2002	(Rel. 21, Last sequence update)			FT CHAIN	65
DE	Cathepsin D precursor	(EC 3.4.23.5).			FT CHAIN	165
CN	Rattus norvegicus (Rat)				FT CHAIN	65
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Rattus. NCBI_TaxID=10116;				FT CHAIN	118
OX	[1]				FT ACT SITE	407
RP	SEQUENCE FROM N.A.				FT ACT SITE	97
RC	SEQUIN-Sprague-Dawley; TISSUE=Pituitary; MEDLINE=91057150; PubMed=2413802;				FT DISULFID	290
RA	Birch N.P.; Loh Y.P.;				FT DISULFID	91
RL	"Cloning, sequence and expression of rat cathepsin D.";				FT DISULFID	110
RN	18:6445-6445(1990).				FT DISULFID	281
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 65-74; 118-127 AND 165-174.				FT CARBOHYD	324
RC	TISSUE-Liver; MEDLINE=1135249; PubMed=3183350;				FT CARBOHYD	134
RA	Fujita H., Tanaka Y., Kono A., Himento M., Kato K.;				FT CARBOHYD	258
RT	"Isolation and sequencing of a cDNA clone encoding rat liver lysosomal cathepsin D and the structure of three forms of mature enzymes.";				FT CONFLICT	15
RT	Biochem. Biophys. Res. Commun. 179:190-196(1991).				FT CONFLICT	163
RL	SEQUENCE OF 134-170. MEDLINE=99034127; PubMed=3182800;				FT CONFLICT	205
RA	Yonemura S., Takahashi T., Wang X., Wong R.N.S., Hartsuck J.A., Tang J.;				FT CONFLICT	262
RT	"Structures at the proteolytic processing region of cathepsin D.";				FT CONFLICT	407 AA;
RL	J. Biol. Chem. 263:1604-1611(1988).				SEQUENCE	44680 MW;
CC	-1- FUNCTION: Acid protease active in intracellular protein breakdown.				SEQUENCE	CA23AD104D5F84 CRC64;
CC	-1- CATALYTIC ACTIVITY: Specificity similar to, but narrower than, chain of insulin. Does not cleave the 4-Gln-1-His-5 bond in B				Query Match	11.1%
CC	-1- SUBUNIT: OCCURS AS A MIXTURE OF BOTH A SINGLE CHAIN FORM AND TWO TYPES OF TWO CHAIN (LIGHT AND HEAVY) FORMS.				Best Local Similarity	26.1%
CC	-1- SUBCELLULAR LOCATION: Lysosomal.				Mismatches	2/3e-16;
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.				Indels	88;
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial purposes requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).				Gaps	20;
DR	EMBL; X51467; CA3349.1;					
DR	PIR; S13111; KRHTD.					
DR	PIR; J01177; J01177.					
DR	HSSP; P07339; LYB.					
DR	MEROPS; A01_009;					
DR	InterPro; IPR01461; AspProteaseAl.					
DR	InterPro; IPR01963; AspProteaseSite.					
DR	Pfam; PF00026; asp_1.					
DR	PRINTS; PR00792; PEPSIN.					
DR	PROSITE; PS00141; ASP_PROTEASE_2.					

Search completed: June 5, 2003, 16:12:44  
Job time : 24 secs



Copyright Gencore version 5.1.6  
(c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 16:11:35 ; Search time 43 Seconds  
(without alignments)  
1120.078 Million cell updates/sec

Title: US-09-869-414A-4  
Perfect score: 2664  
Sequence: 1 MAQALPWLWMAGVLPAH.....CLRCURQHDDFADDISLKK 501

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

#### 8. Summary

Result No. Score Query Match Length DB ID Description

Result No.	Score	Query	Match	Length	DB	ID	Description
1	2656	99.7	501	2	A59090		aspartic proteinas
2	324	12.2	384	2	JC7574		pepsinogen A - Afr
3	313.5	11.8	385	2	JC7575		pepsinogen A - bul
4	309	11.6	387	2	B38302		pepsin (EC 3.4.23.
5	307.5	11.5	383	2	JC7573		pepsinogen C - Afr
6	307	11.5	388	1	S19682		pepsin A (EC 3.4.2
7	305	11.4	384	2	A39314		gastricsin (EC 3.4
8	304	11.4	382	1	PECH		pepsin A (EC 3.4.2
9	301.5	11.3	383	2	A41443		pepsin (EC 3.4.23.
10	301.5	11.3	396	2	A34401		cathepsin E (EC 3.
11	300	11.3	412	1	KHUD		cathepsin D (EC 3.
12	299	11.2	387	2	D38302		pepsin (EC 3.4.23.
13	298	11.2	387	2	D38302		pepsin (EC 3.4.23.
14	297	11.1	407	1	KHRTD		cathepsin D (EC 3.
15	295	11.1	391	2	A43356		cathepsin E (EC 3.
16	294.5	11.1	388	1	S19684		pepsin A (EC 3.4.2
17	290.5	10.9	380	2	I47176		chymosin (EC 3.4.2
18	289	10.8	387	2	E38302		pepsin (EC 3.4.23.
19	288.5	10.8	388	1	PEHU		pepsin A (EC 3.4.2
20	288.5	10.8	388	1	PEMQAR		pepsin A (EC 3.4.2
21	288.5	10.8	388	2	A30142		pepsin A (EC 3.4.2
22	288.5	10.8	398	2	S66465		cathepsin E (EC 3.
23	288	10.8	389	2	JB0371		pepsin C (EC 3.4.2
24	287.5	10.8	444	2	T24204		hypothetical prote
25	287	10.8	410	1	KHMSD		cathepsin D (EC 3.
26	286.5	10.8	388	1	PEMQAJ		pepsin A (EC 3.4.2
27	286	10.7	398	2	I51185		cathepsin D (EC 3.
28	285.5	10.7	388	2	B30142		pepsin A (EC 3.4.2
29	284.5	10.7	381	1	CMSHB		chymosin (EC 3.4.2

#### Alignments

#### RESULT 1

A59090

aspartic proteinas (EC 3.4.23.-) BACE precursor - human

N;Alternate names: beta-secretase; beta-site APP cleaving enzyme

C:Species: Homo sapiens (man)

C;Date: 2-Oct-1999 #sequence-revision 29-Oct-1999 #text\_change 11-May-2000

C;Accession: A59090

R;Vassar, R.; Bennett, B.D.; Babu-Khan, S.; Mendiaz, E.A.; Denis, P.; Teplo

M.A.; Biere, A.L.; Curran, E.; Burgess, T.; Louis, J.C.; Collins, F.; Treanor, J.; Ro

Science 286, 735-741, 1999

A;Title: beta-Secretase cleavage of Alzheimer's amyloid precursor protein by the tran

A;Reference number: A59090; PMID:20002972; PMID:10531052

A;Note: Submitted to GenBank, September 1999

A;Accession: A59090

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Cross-references: GB:AF190725; NID:96118538; PLDN:AAF04142.1; PID:96118539

C;Genetics:

A;Gene: BACE

C;Superfamily: beta-secretase

C;Keywords: Alzheimer's disease; aspartic proteinas; brain; glycoprotein; hydrolase;

F;1-21/Domain: signal sequence #status predicted <SIG>

F;22-45/Domain: propeptide #status predicted <PRO>

F;46-501/Domain: acid proteinase BACE #status predicted <ATP>

F;461-477/Domain: transmembrane #status predicted <TRN>

F;93-289/Active site: Asp #status predicted

F;153-172-223-354/Binding site: carbohydrate (Asn) (covalent) #status predicted

F;330-380/Disulfide bonds: #status predicted

Query Match 99.7%; Score 2656; DB 2; Length 501;

Best Local Similarity 99.8%; Pred. No. 7e-206;

Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Y MAQALPWLLWMMAGVLPWAHGTOHGTQHTRLPLRSGLPGLQLPRTDEEPEPGRGSF 60

1 MAQALPWLLWMMAGVLPWAHGTOHGTQHTRLPLRSGLPGLQLPRTDEEPEPGRGSF 60

Y MAQALPWLLWMMAGVLPWAHGTOHGTQHTRLPLRSGLPGLQLPRTDEEPEPGRGSF 60

1 YRDLRKGVYPTQKWEGLGTDVSIPHPVNTVRNATAITESDKFTTNGSWEGIL 180

121 YRDLRKGVYPTQKWEGLGTDVSIPHPVNTVRNATAITESDKFTTNGSWEGIL 180

Y 61 VEMVDNLRLRGSGQYYVEMTVGSPPTQLNLIVDGSNNFANGAAFPFLHRYYQRLSS 120

61 VEMVDNLRLRGSGQYYVEMTVGSPPTQLNLIVDGSNNFANGAAFPFLHRYYQRLSS 120

Y 121 GLAYEAIARPDDSIPFFEDSLVKOTHVNPNUFLHCGAGPQLNQSEVLASYGGSMIGGI 240

181 GLAYEAIARPDDSIPFFEDSLVKOTHVNPNUFLHCGAGPQLNQSEVLASYGGSMIGGI 240

181 GLAYEAIARPDDSIPFFEDSLVKOTHVNPNUFLHCGAGPQLNQSEVLASYGGSMIGGI 240

241 DHSLTGSLWYTPIRREWYEVIVTRVETINGQDLKMDCKEYNDSIVDGTNULRPLKK 300

1 CMSHB



A; Title: Structure and development of rabbit pepsinogens. Stage-specific zymogens, nucleic acids, and mRNA. PMID:2129536

A; Reference number: A38302; MUID:9109127

A; Accession: B38302

A; Status: preliminary

A; Molecule type: mRNA

A; Residues: 1-387 <KAG>

A; Cross-references: GB:M59235; GB:J05638

C; Superfamily: Pepsin

C; Keywords: aspartic proteinase; hydrolase; phosphoprotein; protein digestion

Query Match Score 309, DB 2; Length 387;

Best Local Similarity 27.1%; Pred. No. 5.1e-17;

Matches 98; Conservative 68; Mismatches 130; Indels 66; Gaps 15;

Db 7 YVEMTVGSPPQTNLILVDTGSNNPAVG ---AAPHFPLHRYQFQLSSTYRDLRKGVVV 130

Db 75 YFGTISGTPPQEFTVIFDTGSNNLWVPTSYCSSLACFLHKRENDDSTFQATSETLSI 134

Db 131 PYTQGRKGWEGLGTDLVSIHPGPNTVRAIAITESD - KFFINGSNNEGIGLGLAYAEI 187

Db 135 TYGTGSTMGIQGYDVKV --GNIDDTNQFLGSKIEPQTFLY -AAPPFDGLGLAYPSI 189

Db 188 ARPDDSLPEPFDLSLVKOTHIV -PNLFSLHLGAGFPFLNQSEVLASVGGSMILLGIDHSLYT 246

Db 190 SASDAT--PVEDNMNEGLYSEDIFSVYLLSSNG-----ERGSMMNGGFDSSYTT 237

Db 247 GSLWTPPIRREWWYEVIVRVEINQDLMK - DCKBYNYDKSIVOSGTINLRLPKVFEA 304

Db 238 GSLNNWPVSHEGYWQITMDSTTINGETIACADSC ---QAVVDTGTSLLAGPTSAISK 291

Db 305 AVKSKAAASSTEKFDPDFWIGEQLV-CWQAGTTPWNIFPVTSVLYMGEVTNQSFRITILP 363

Db 292 IGYIGASKNL-----LGENIISCSAIDSLSLPDIFV -----TINN 325

Db 364 QYLREVED-VATSQQDC - YKFAISQSSSTGT - VMGAVIMEGTYVVFDRARKRIGFAV 417

Db 326 VQPLPASAVLKEEDDCCLSGFDGMNLDTSYGELMLGDYFIROQFTVFDRANNQYGLAA 385

Db 418 SA 419

Db 386 AA 387

RESULT 5

JC7573

pepsinogen C - African clawed frog

N; Alternate names: progastrins

Species: Xenopus laevis (African clawed frog)

C; Accession: JC7573; PC7118

R; Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasunusu, S.

J. Biochem. 129, 147-153, 2001

A; Reference number: JCT573; MUID:21064922; PMID:11134569

A; Accession: PC7118

A; Molecule type: protein

A; Residues: 1-383 <IKU>

A; Cross-references: DDBJ:AB045379

A; Accession: PC7118

A; Molecule type: protein

A; Residues: 17-68 <IK2>

C; Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-like a

C; Generics:

C; Superfamily: pepsin

C; Keywords: zymogen

Query Match Score 307.5; DB 2; Length 383;

Best Local Similarity 25.9%; Pred. No. 6.6e-17;

Matches 112; Conservative 64; Mismatches 139; Indels 117; Gaps 19;

Db 23 QHGIRLPLRSGLGGAPLGRLPRETDEEEPPGRGSFVMVDLRLRGSKSCQGYVEMTVG 82

Db 34 ENGIAPL-----VDPATKVVNOVATAYEPLSNMMDMS--YYGEISIG 74

Db 83 SPPOTINILVDTGSSNFNAYGA-----APHPFELHRYIQRLSSTYRDLRKGVVYPYQ 134

Db 75 TPPQNFLVLFDTGSNLWVASTYCCSQACTNHNPL---FNPQSOSTSSNQQOFSIQLQGT 130

Db 135 GKWEGEGLDLVSIHPGPNTVRAIAITESDKEFFINGN-----WEGIUGLAVAEI 189

Db 131 GSLTGGLGYDVVTI ---QVNAISQEQFLSETEP---GYNFVYQFDGTLGLAPSIA 182

Db 190 PDDSLPEPFEDSLVKOTHVPN-LFSLSHLGAGFPFLNQSEVLASVGGSMIIIGGIDHSLYTGS 248

Db 183 -VGGATTVMQGMMQNLNQNPQFGLYSG----QS---SQNGGEVAFGVGDQNYYTQG 232

Db 249 LWYTPFIREWYEVIVRVEINQDQD - LAMDCKEYNDNSIVDGTSGTNRURLPKVFEAA 305

Db 233 IYWPVPTSETYWOIQLQGFSINGQATGWCNSOGC----QAVIDGTGTSLLTAPOQSVSSL 286

Db 306 VKSTKAAASSTEKFDPDFWIGEQLVQWQAGTIPWNI -FPVLSLYIMG-----EVTN 354

Db 287 IQSIGAQQDON-----GQYVVSQS-----NQNLPITSFTISGVSPPLPPSAYVHQ 332

Db 355 QS-----FRUTILPQQYLRPVEDAVTSODDCYKFAISOSSTGTVMGAVIMEGRYVVF 407

Db 333 QSSGQCTIGIMPYLPSONGQPL-----WILGDVFLRETVSYVD 371

Db 408 RAKRIGFAVSA 419

Db 372 LGNNQVGFATAA 383

RESULT 6

S19682

Pepsin A (EC 3.4.23.1) 4 precursor - Japanese macaque

N; Alternative names: pepsinogen A isozyme 4

C; Species: Macaca fuscata (Japanese macaque)

C; Date: 22-Nov-1993 #sequence\_revision 19-Oct-1995 #text\_change 18-Jun-1999

C; Accession: S19682; S1605

R; Kageyama, T.; Tanabe, K.; Koivali, O.

Eur. J. Biochem. 202, 205-215, 1991

A; Title: Development-dependent expression of isozymogens of monkey pepsinogens and st

A; Reference number: S19681; PMID:1935977

A; Accession: S19682

A; Molecule type: mRNA

A; Residues: 1-388 <TAG>

A; Cross-references: EMBL:X59753; PID:938071

A; Note: parts of sequence, including amino ends of pepsinogen and activation intermed

C; Comment: This is a minor component of pepsin at all post-partum stages.

C; Superfamily: pepsin

C; Keywords: aspartic proteinase; gastric juice; hydrolase; phosphoprotein; protein di

F; 1-15/Domain: signal sequence #status predicted <SIG>

F; 16-388/Product: pepsinogen A #status experimental <PPR>

F; 63-#Status: experimental <ENZ>

F; 38-39/Cleavage site: Leu-Lys (pepsin) #status experimental

F; 62-63/Active site: Leu-Lys (pepsin) #status experimental

F; 94-277/Active site: Asp #status predicted

F; 107-112,268-272,311-314/Disulfide bonds: #status predicted

F; 130/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match Score 307.5; DB 1; Length 388;

Best Local Similarity 27.8%; Pred. No. 6.8e-17;

Matches 108; Conservative 65; Mismatches 135; Indels 83; Gaps 17;

Db 44 PRETDEEEPPGRGSFVMDNLRGSGQGYVYMTVGSPPOTTINLYDTGSSNFAVG 103

Db 60 PTLDQPLE ---NYLDV -----BYFGTIGIGPAQNFTVEDGSSLWV - 102

Db 104 APHPFL-----HYYQRLSSTYRDLRKGVVYPYQTKWEGELGTDLVSTPHGPNTVY 156

Db 103 -PSVYCSSLACMDHNUFNPOQDSSTYRATSKTVSTYGTGSMTGILGYDTVKV - GGISD 158



C;Accession: A41443  
R; Hayashi, K.; Agata, K.; Mochii, M.; Yassugi, S.; Eguchi, G.; Mizuno, T.  
J. Biochem. 103, 290-296, 1988  
A;Title: Molecular cloning and the nucleotide sequence of cDNA for embryonic chicken **pepC**  
A;Reference number: A41443; MUID:88227903; PMID:3131317  
A;Accession: A41443  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-383 <AZB>  
A;Cross references: GB:D00215; NID:92760810; PIDN:BAA00153.1; PID:9222853  
C;Superfamily: pepsin  
C;Keywords: aspartic proteinase; hydrolase; protein digestion

Query Match 11.3% Score 301.5; DB 2; Length 383;  
Best Local Similarity 25.2%; Pred. No. 2e-16;  
Matches 90; Conservative 76; Mismatches 124; Indels 67; Gaps 14;

<p>Qy 75 YYVEMTVSSPPQTPLNLINLWDTGSSNFAYGA---APHPEFLHRYYQRLSLSITYDLRKGVYY 130  <b>76 YYGTISITGSPQDFETVVFDTGSSNLWVFSVCTSPAQSHQFNPSQSSTIKSTGQNLSI 135</b></p> <p>2 131 PYTQKGKNGELGLDLYSPHGPNTVANIAAATESDKPFTTGSNNWKGIGLGLAYAETARP 190  <b>136 HYTGDMGCTVGTDVTVASLMDTNQLGLSLT-SEPGCFEVY-VKFDGLGIGPSLIA-</b> 192</p> <p>Qy 191 DDSLEPFEDSLVLYKOTHV-PNLFSLHLICGAGEPLINQSEVYLASYGSGMTIGIDHSLYTGS 249  <b>193 -DGTPVDTDNMYNESLLEQNLFSVYL-----REPGSMVYFGGLIDESYFTGSI 240</b></p> <p>Db 250 WTPPIRRWYEVITVRVEINGDL---KMDCKEYNDSIVDGTGTMNLRPKVFEAVK 307  Qy 241 NWIPVSYQGWQISMDSLIVNQEIACSSGC-----QAIIDTGTSLVAGPASDINDQS 294</p> <p>Qy 308 SIKAASSTKEKEPGEWILCEOLQVQCWAQCTTPWNTEFPVSL-----YLMCEBTVNGSFRLTLP 363  Db 205 AVGANQNNT-----YGEYSV-----NCISHILLAMPDVVFVIGI----- 326</p> <p>Qy 364 QQYIRPVDEVA----TSQDDCYKFAISSSSTGTVMGAVIMEGFYVVFDARKRIGFA 416  Db 327 -QY---PVPALAYTEQNGGCTCMSSFQNNSADLWILGDVFIRVYYSIDRANNRVGLA 380</p>	<p>R; Hill, J.; Montgomery, D.S.; Kay, J.  BBS Lett. 326, 101-104, 1993  A;Title: Human cathepsin E produced in <i>E. coli</i>.  A;Reference number: S34467; MUID:93314762; PMID:8325357  A;Accession: A41447  A;Status: preliminary  A;Molecule type: protein  A;Residues: 57-60, 62-61 &lt;AZD&gt;  R; Athauda, S.B.P.; Matsushita, O.; Kageyama, T.; Takahashi, K.  Biochem. Biophys. Res. Commun. 168, 87-885, 1990  A;Title: Structural evidence for two isoformic forms and the carbohydrate attachment s  A;Reference number: A34643; MUID:90241267; PMID:2334440  A;Accession: A34643  A;Status: preliminary  A;Molecule type: protein  A;Residues: 54-58, 'XXX', 62-64, 'M', 66-89, 'X', 91-95 &lt;ATH&gt;  A;Cross references: GDB:119821; OMIM:116890  A;Map position: 1q31-1q31  C;Superfamily: pepsin  C;Keywords: aspartic proteinase; blocked amino end; hydrolase; zymogen  F; 1-17/Domain: signal sequence #status predicted &lt;SIG&gt;  F; 18-53/Domain: activation peptide #status predicted &lt;PRO&gt;  F; 54-396/Product: cathepsin E #status predicted &lt;MAT&gt;  F; 18/Modified site: blocked amino end (Gln) (in mature form) (probably pyrrolidone ca  F; 96, 281/Active site: Asp #status predicted</p> <p>Query Match 11.3%; Score 301.5; DB 2; Length 396;  Best Local Similarity 25.8%; Pred. No. 2.1e-16;  Matches 100; Conservative 68; Mismatches 144; Indels 75; Gaps 16;</p> <table border="0"> <tr> <td style="vertical-align: top;"> <p>Qy 48 DEEPEEPGRGSFEMYMDNLRGKSQGYVYEMTVGSPQPTLNLTVDPGSSNEFAVGA--- 103  Db 63 DQSAKEP-----LINYLD-----MEVFGTISGQNFVTFQDGSNNLWPSVYCT 110</p> <p>Qy 104 APHPFLHRYYQRLSSTYDLRKGVYYVPTQKGWEGLDLYSPHGPNTVRAIAI 163  Db 111 SPACKTHSRFOPSSSTYSGQGOFSTIOTYGTGSLSLIGADQVS-EGLTVYQGQFESV 169</p> <p>Qy 164 TESDKFFINGSWBEGIQLGLAYAEIARPDDSLEPFDSLVYKQHPV-P-NFLSLHLUGAGGPL 222  Db 170 TEQGTFED-AEFDGLLGIGPSLIA-VGGVTPFEDNMMAQNLVDLPMFSVYM----- 219</p> <p>Qy 223 NQSEVLSVGGSMTIGGIDHSLYTGSILWYTPIRRENYEVILVRENGODLKMDCKEYN 282  Db 220 -SSMPSEGAGSELTFGVDHSHFGSLNWVPTQAYQIALDNIQVG---TVMFCSB-- 274</p> <p>Qy 283 YDKSIVDSGTTNULRPLPKVFEAKVSIKAASSTEKFDPGEWLGEQLVQWQAGTTPNWIFP 342  Db 275 GCOAIVDGTSLITGSPDKIKLQNAIGAAP-----CANLNVMP 321</p> <p>Qy 343 VISLYLMEGVNTOSFRITLPOOLYLPVDEVATSDQDCYKFAISQSSTG----- 391  Db 322 DVTFTING-----VPYTLSPSTAY-TLIDFVDDMQFC-----SSGFQGLDIHPAG 365</p> <p>Qy 392 --TVMGAVIMEGFWFDRKRGFA 416  Db 366 PLWILGDVFIRQFYSVFDRGNNEVGLA 392</p> </td> </tr> </table> <p>RESULT 10</p> <p>A34401 cathepsin E (EC 3.4.23.34) precursor - human  C;Species: Homo sapiens (man)  C;Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 22-Jun-1999  C;Accession: A42038; M34401; S35663; S34463; B31643  R;Zuma, T.; Liu, W.; Vander Laan, D.J.; Bowcock, A.M.; Taggart, R.T.  J. Biol. Chem. 267, 1609-1614, 1992  A;Title: Human gastric cathepsin E. Predicted sequence localization to chromosome 1, and  A;Reference number: A42038; MUID:1370478  A;Accession: A42038  A;Molecule type: DNA  A;Residues: 1-396 &lt;AZD&gt;  A;Cross references: GB:J05036; NID:9181193; PIDN:AAA52130.1; PID:g181194  R;Takeda-Enzaki, M.; Yamamoto, K.  Arch. Biochem. Biophys. 304, 352-358, 1993  A;Title: Isolation and biochemical characterization of procathepsin E from human erythrocytes  A;Reference number: S35663; MUID:93349047; PMID:8344912  A;Accession: A414401  A;Molecule type: mRNA  A;Residues: 1-396 &lt;AZB&gt;  A;Cross references: GB:J05036; NID:9181193; PIDN:AAA52130.1; PID:g181194  R;Faust, P.L.; Kornfeld, S.; Chirgwin, J.M.</p> <p>RESULT 11.</p> <p>KHHUD cathepsin D (EC 3.4.23.5) precursor [validated] - human  N;Alternative names: proprocathepsin D  C;Species: Homo sapiens (man)  C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 15-Sep-2000  C;Accession: A25711; S30749; PC066; 159246; 157716</p>	<p>Qy 48 DEEPEEPGRGSFEMYMDNLRGKSQGYVYEMTVGSPQPTLNLTVDPGSSNEFAVGA--- 103  Db 63 DQSAKEP-----LINYLD-----MEVFGTISGQNFVTFQDGSNNLWPSVYCT 110</p> <p>Qy 104 APHPFLHRYYQRLSSTYDLRKGVYYVPTQKGWEGLDLYSPHGPNTVRAIAI 163  Db 111 SPACKTHSRFOPSSSTYSGQGOFSTIOTYGTGSLSLIGADQVS-EGLTVYQGQFESV 169</p> <p>Qy 164 TESDKFFINGSWBEGIQLGLAYAEIARPDDSLEPFDSLVYKQHPV-P-NFLSLHLUGAGGPL 222  Db 170 TEQGTFED-AEFDGLLGIGPSLIA-VGGVTPFEDNMMAQNLVDLPMFSVYM----- 219</p> <p>Qy 223 NQSEVLSVGGSMTIGGIDHSLYTGSILWYTPIRRENYEVILVRENGODLKMDCKEYN 282  Db 220 -SSMPSEGAGSELTFGVDHSHFGSLNWVPTQAYQIALDNIQVG---TVMFCSB-- 274</p> <p>Qy 283 YDKSIVDSGTTNULRPLPKVFEAKVSIKAASSTEKFDPGEWLGEQLVQWQAGTTPNWIFP 342  Db 275 GCOAIVDGTSLITGSPDKIKLQNAIGAAP-----CANLNVMP 321</p> <p>Qy 343 VISLYLMEGVNTOSFRITLPOOLYLPVDEVATSDQDCYKFAISQSSTG----- 391  Db 322 DVTFTING-----VPYTLSPSTAY-TLIDFVDDMQFC-----SSGFQGLDIHPAG 365</p> <p>Qy 392 --TVMGAVIMEGFWFDRKRGFA 416  Db 366 PLWILGDVFIRQFYSVFDRGNNEVGLA 392</p>
<p>Qy 48 DEEPEEPGRGSFEMYMDNLRGKSQGYVYEMTVGSPQPTLNLTVDPGSSNEFAVGA--- 103  Db 63 DQSAKEP-----LINYLD-----MEVFGTISGQNFVTFQDGSNNLWPSVYCT 110</p> <p>Qy 104 APHPFLHRYYQRLSSTYDLRKGVYYVPTQKGWEGLDLYSPHGPNTVRAIAI 163  Db 111 SPACKTHSRFOPSSSTYSGQGOFSTIOTYGTGSLSLIGADQVS-EGLTVYQGQFESV 169</p> <p>Qy 164 TESDKFFINGSWBEGIQLGLAYAEIARPDDSLEPFDSLVYKQHPV-P-NFLSLHLUGAGGPL 222  Db 170 TEQGTFED-AEFDGLLGIGPSLIA-VGGVTPFEDNMMAQNLVDLPMFSVYM----- 219</p> <p>Qy 223 NQSEVLSVGGSMTIGGIDHSLYTGSILWYTPIRRENYEVILVRENGODLKMDCKEYN 282  Db 220 -SSMPSEGAGSELTFGVDHSHFGSLNWVPTQAYQIALDNIQVG---TVMFCSB-- 274</p> <p>Qy 283 YDKSIVDSGTTNULRPLPKVFEAKVSIKAASSTEKFDPGEWLGEQLVQWQAGTTPNWIFP 342  Db 275 GCOAIVDGTSLITGSPDKIKLQNAIGAAP-----CANLNVMP 321</p> <p>Qy 343 VISLYLMEGVNTOSFRITLPOOLYLPVDEVATSDQDCYKFAISQSSTG----- 391  Db 322 DVTFTING-----VPYTLSPSTAY-TLIDFVDDMQFC-----SSGFQGLDIHPAG 365</p> <p>Qy 392 --TVMGAVIMEGFWFDRKRGFA 416  Db 366 PLWILGDVFIRQFYSVFDRGNNEVGLA 392</p>		

Proc. Natl. Acad. Sci. U.S.A. 82, 4910-4914, 1985  
 A;Title: Cloning and sequence analysis of cDNA for human cathepsin D.  
 A;Reference number: A25771; MUID:3927292  
 A;Molecule type: mRNA  
 A;Residues: 1-412 <FAUD>  
 A;Cross-references: EMBL:MI1233; NID:9181179; PIDN:AAB59529.1; PMID:9181180  
 Nucleic Acids Res. 15, 3773-3776, 1987  
 A;Title: Oestrogen responsive cathepsin D mRNA levels in oestrogen responsive human breast  
 A;Reference number: S30749; MUID:87231068; PMID:3588310  
 A;Accession: S30749  
 A;Molecule type: mRNA  
 A;Residues: 1-412 <WES>  
 A;Cross-references: EMBL:X05341; NID:929677; PIDN:CAA28955.1; PMID:929678  
 R;May, F.E.B.; Smith, D.J.; Westley, B.R.  
 Gene 134, 277-282, 1993  
 A;Title: The human cathepsin D-encoding gene is transcribed from an estrogen-regulated a  
 A;Reference number: PC2066; MUID:94085791; PMID:822386  
 A;Accession: PC2066  
 A;Molecule type: DNA  
 A;Residues: 1-23 <MAV>  
 A;Cross-references: GB:L12980; NID:9291930; PIDN:AAA16314.1; PMID:9455429  
 A;Experimental source: MCF-7 cell  
 R;Cavailles, V.; Augereau, P.; Rochefort, H.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 203-207, 1993  
 A;Title: Cathepsin D gene is controlled by a mixed promoter, and estrogens stimulate on  
 A;Reference number: I59236; MUID:9312342; PMID:8419924  
 A;Accession: I59236  
 A;Status: translation not shown; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-22 <CAV1>  
 A;Cross-references: GB:552557; NID:9263124; PIDN:AA013868.1; PMID:94261568  
 R;Augereau, P.; Miralles, F.; Cavailles, V.; Gadellet, C.; Parker, M.; Rochelefort, H.  
 Mol. Endocrinol. 8, 693-703, 1994  
 A;Title: Characterization of the proximal estrogen-responsive element of human cathepsin  
 A;Reference number: I57716; MUID:95021301; PMID:7935485  
 A;Accession: I57716  
 A;Status: translation not shown; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-22 <CAV2>  
 A;Cross-references: GB:S74689; NID:978630; PIDN:AA014156.1; PMID:94261856  
 R;Baldwin, E.T.; Bhat, T.N.; Gulinik, S.; Erickson, J.W.  
 Submitted to the Brookhaven Protein Data Bank, April 1993  
 A;Reference number: A51839; PDB:1IYA  
 A;Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 65-161;170-241  
 R;Baldwin, E.T.; Bhat, T.N.; Gulinik, S.; Erickson, J.W.  
 Submitted to the Brookhaven Protein Data Bank, April 1993  
 A;Reference number: A51840; PDB:1IYB  
 R;Baldwin, E.T.; Bhat, T.N.; Gulinik, S.; Hosur, M.V.; Sowder, II, R.C.; Cachau, R.E.; Col  
 Proc. Natl. Acad. Sci. U.S.A. 90, 6796-6800, 1993  
 A;Title: Crystal structures of native and inhibited forms of human cathepsin D: implicat  
 A;Reference number: A48229; MUID:9334306; PMID:8393577  
 A;Contents: annotation; X-ray crystallography, 2.5 angstroms  
 C;Comment: Cathepsin D is a ubiquitous lysosomal proteinase.  
 C;Comment: In addition to the Propeptide residues 163-168 and 411-412 are proteolyticall  
 C;Comment: The carbohydrate bound to 134-Ash contains a mannose-6-phosphate that is bound  
 A;Gene: GDB:CTSD  
 A;Map position: 11p15.5-11p15.5  
 C;Function:  
 A;Cross references: Limited specificity endopeptidase  
 A;Pathway: intracellular protein degradation  
 C;Superfamily: pepsin  
 C;Keywords: aspartic proteinase; glycoprotein; hydrolase; lysosome; protein degradation  
 F:1-20/Domain: signal sequence #status predicted <SIG>  
 F:21-64/Domain: propeptide #status predicted <PRO>  
 F:65-164/Domain: cathepsin D #status experimental <MAT>  
 F:267,329-356/Region: phosphotransferase recognition  
 F:91-160,110-117,286-290,329-366/Disulfide bonds: #status experimental  
 F:97,295/Active site: Asp #status experimental

F;134,263/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 Query Match 11.3%; Score 300.5; DB 1; Length 412;  
 Best Local Similarity 26.9%; Pred. No. 2.7e-16;  
 Matches 123; Conservative 68; Mismatches 1.0%; Indels 97; Gaps 21;  
 QY 5 LPWLILWMGAGVLPAGTQHGLPLR-----SSLGGAPLGL-----RLP 44  
 Db 7 LPIALCLIAA--PASAL--VRLPLKETSIIRTMSVPGSVDIATGPVSKSYDAPV 60  
 QY 45 RETDIEPEEPGRROSSFVENVNLRKGSSCGYYVEMTGVSSPPQTLNLVDTGSNFNAVGA 104  
 Db 61 AVTE -----GPIPEVKNYMDAQ---YYEGIGTTPPPQCFVVFDTGSSNLWVPSI 108  
 QY 105 PHPFL-----HRYYQRLSSTYRDLRKGVVYVPTQGRWEGEGLTDLVSLIP-----149  
 Db 109 HCKLDLADIACWIMHKYNSDKSSTYVNGTSFDIHLYSGLSLQDGTYSVPOSASSASA 168  
 QY 150 HGPNTVTRANIAATEESDKFFINGSNWEGTSLGLAYAEIARPDDSLVLYKOTHY-P 208  
 Db 169 LGGVKVERQVGEAKTFIARAKFDGLGMAYPRLS-VNNVLPPFDNLMQKLVDQ 226  
 QY 209 NLFSLHLCCGAGFPLNQSEVLAvgGSMITGGIDHSLYTGSFLWYTPIRREWWYEVLYRVE 268  
 Db 227 NIFSYL-----SRDPDAQGCFMLGGTDSKYYKCSLSLTNVTRKAYWQHDLQVE 278  
 QY 269 T-NGQDLKMDCKEYNNVDSKLVDSGTNTNLRLPKVYEAVKSIKAASSTEKEPDGFWLGEQ 327  
 Db 279 VASGLTL--CKE--GCEALVDTGTSLMLYGPVDEVERLOKATGAVPLIO-----325  
 QY 328 LV-CWQAGTTPWNIFPVISLYLMGVIVINOSERFILIPQQLRVPEDVATSQDDCYKFAT 386  
 Db 326 MIPCEKVST---LPATILKLGG---KGYKLS-PEDYTLKVSQAGKTL-CLSGPMG 372  
 QY 387 Q----SSTGTGTMGAVMEGFYVVFDRARKRIGFEAVSA 419  
 Db 373 MDIPPPSGPLWILGDVTGTRYTFDRDNRNVRGFAEEA 410

RESULT 12  
 C38302 pepsin (EC 3.4.23.-) II-2/3 precursor - rabbit  
 C;Species: Oryctolagus cuniculus (domestic rabbit)  
 C;Date: 14-Jun-1991 #sequence\_revision 14-Jun-1991 #text\_change 23-Feb-1997  
 C;Accession: C38302  
 R;Kageyama, T.; Tanabe, K.; Kawai, O.  
 J. Biol. Chem. 265, 17031-17038, 1990  
 A;Title: Structure and development of rabbit pepsinogens. Stage-specific zymogens, nu  
 A;Reference number: A38302; MUID:91009127; PMID:2129536  
 A;Accession: C38302  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-387 <KAG>  
 A;Cross references: GB:J05638  
 C;Superfamily: pepsin  
 C;Keywords: aspartic proteinase; hydrolase; phosphoprotein; protein digestion  
 Query Match 11.2%; Score 299; DB 2; Length 387;  
 Best Local Similarity 26.9%; Pred. No. 3.3e-16;  
 Matches 97; Conservative 64; Mismatches 134; Indels 66; Gaps 13;

QY 75 YYVEMTVGSPPOTLNLVDTGSNFNAVGAHPF-----LHRYYQRQLSSTYRDLRKG 127  
 Db 75 YFGTISIGTPQDDETFVDTGSSNLWV--PSTYCSSLACALHKRKNPEDSSTYQCTSET 131  
 QY 128 VVBYPTQKWEGEGLDLYSIPHGPNPNTVRNATAITESDKFFINGSNWEGLGLAYAEI 187  
 Db 132 LSITYGTGSMNTGTYDPTVKGSEDINQIFGLSKTEPSLTLF-APFDGILGLAYPSI 189  
 QY 188 ARPDDSLPFFDSLVLYKOTHY-PNLFSLHLCGAGFPLNSEVLASG 246  
 Db 190 SSSDAT--PVFDNNWNNEGQLVSDQFLSVYLSSDD-----EKGSLLVMFOGTTDSSYYT 237



Db 330 ---LPIITPKLGQ-----NYELHPEYILKVSQAGRT-----ICISGFNGMDIIPP 373  
 Qy 388 SSTGTVMGAVIMEGFVVFDEARKRIGFAVSA 419  
 Db 374 SGPLWILGDVFFIGCYTVDREYNVGFKA 405

## RESULT 15

A43356 cathepsin E (EC 3.4.23.34) precursor - guinea pig  
 N;Alternate names: erythrocyte membrane aspartic proteinase; slow-moving proteinase  
 C;Species: Cavia porcellus (guinea pig)  
 C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
 C;Accession: A43356  
 R;Kageyama, T.; Ichinose, M.; Tsukada, S.; Miki, K.; Kurokawa, K.; Koiwai, O.; Tanji, M.  
 J. Biol. Chem. 267, 16450-16459, 1992  
 A;Title: Gastric procathepsin E and progastricsin from guinea pig. Purification, molecular  
 A;Reference number: A43356; MUID: 92355614; PMID:1644829  
 A;Accession: A43356  
 A;Molecule type: mRNA  
 A;Residues: 1-391 <KAG>  
 A;Cross-references: GID:MM86653; NID:9191294; PID:AAA37052.1; PID:g191295  
 A;Note: sequence extracted from NCBI backbone (NCBIn:110763, NCBI\_P:110769)  
 C;Superfamily: pepsin  
 C;Keywords: aspartic proteinase; hydrolase; membrane protein

Query Match Score 295; DB 2; Length 391;  
 Best Local Similarity 26.98; Pred. No. 7e-16;  
 Matches 98; Conservative 64; Mismatches 130; Indels 72; Gaps 16;

Qy 75 YYVEMTVGSPPQTNLIVDTGSSNFAYGA---APHEFLHRYYOROLSTYRDLRKGVYY 130	Db 74 YFGPTISGQPQQNFTVIDTGSNNLWVFSVCTSPACQTHYPHPHSLSIYSTREVGNFSI 133
Qy 131 PYTGKWEFEGFLDVLSPHGPNVTYRANIAAITESDKFFTSNNNEGTLGHAYABTARP 190	Db 134 QYGTGSLTGIGAQSV-EGLTVVQGOFGESVQEPGKTFVH-AEFDIGLGYPSLA- 190
Qy 191 DDSLEPFDLSVLKOTHVPNLFSLHLCCGFLQPLNOSEVIALASVGSMILGGIHDHSLVYTGSLW 250	Db 191 -GGTPVFDNMMAQ---NLVALPM---FSYMSNSNGSSELTRGYDPDSHFSSSLN 241
Qy 251 YPTTERWWYEVILYRVIEINGQDLKMDKEYNDKSTIVDQCTINRLPKYFEAAYSIK 310	Db 242 WWPYTKQAYWQIALDGIQVW-DSYMECSE--GCQAIWTGSLITGE----PGKIKQLQ 293
Qy 311 AASSTEKFDPDFWLGEOLVQWOAGSTPWNIFPVISLYMGEVTNQSFRI----TILPQQ 365	Db 334 YLRPVEDVATSQDDCYKFAISQSSTG------ANLNMMLDVT--FLLNGVPTLNPTA 333
Qy 394 EALGATYVDEGY----SVQC------ 294	Qy 366 YLRPVEDVATSQDDCYKFAISQSSTG------TVMGAVIMGGFYVVFDRARKR 412
Qy 413 IGFA 416	Db 384 VGLA 387

Db 384 VGLA 387

Search completed: June 5, 2003, 16:17:40  
 Job time : 44 secs

Result No.	Score	Query	Match	Length	DB	ID	Description
1	250.6	Q9BYC1	Q9BYC1	94.1	476	4	Q9BYC1 homo sapien
2	2478.5	Q9BYC1	Q9BYC1	93.0	532	4	Q9BYC1 homo sapien
3	2298	Q9BYC0	Q9BYC0	90.0	457	4	Q9BYC0 homo sapien
4	2285.5	Q9BYB9	Q9BYB9	84.8	432	4	Q9BYB9 homo sapien
5	1407	Q9CUU5	Q9CUU5	52.8	266	11	Q9CUU5 musculus
6	1155.5	Q9H2V8	Q9H2V8	43.4	439	4	Q9H2V8 homo sapien
7	1150	Q9JL18	Q9JL18	43.2	514	11	Q9JL18 musculus
8	974.5	Q9NPL2	Q9NPL2	36.6	468	4	Q9NPL2 homo sapien
9	960.5	Q9NL1	Q9NL1	36.4	396	4	Q9NL1 homo sapien
10	712.5	Q9P0D2	Q9P0D2	26.7	213	4	Q9P0D2 homo sapien
11	596.5	Q9R1P7	Q9R1P7	22.4	255	11	Q9R1P7 musculus
12	354.5	Q9WQY9	Q9WQY9	13.3	244	5	Q9WQY9 aphelinus
13	335.5	Q9YKPF6	Q9YKPF6	12.6	391	5	Q9YKPF6 drosophila
14	335	Q9GYX7	Q9GYX7	12.6	354	3	Q9GYX7 boophilus m
15	319	Q9DEC2	Q9DEC2	12.0	384	13	Q9DEC2 xenopus lae.
16	313.5	Q9DEC4	Q9DEC4	11.8	385	13	Q9DEC4 rana catesbe

us-09-869-414a-4.rsp

Query Match	90.0%	Score 2398; DB 4; Length 457;	Pred. No. 5e-173;	Best Local Similarity	86.2%;	Pred. No. 5e-173;	
Best Local Matches	456; Conservative Matches 432;	Conservative Matches 432;	Mismatches 0;	Local Similarity	86.2%;	Local Similarity 86.2%;	
Qy	1 MAQALPWLLWMAGVLPNHGTQHIGIRLRLPRTDEPEEPGRSSF	Length 457; Pred. No. 3.4e-180; 0; Mismatches 1; Indels 44; Gaps 1;	0; Mismatches 0;	Indels 69;	Gaps 1;	Indels 69; Gaps 1;	
Db	1 MAQALPWLLWMAGVLPNHGTQHIGIRLRLPRTDEPEEPGRSSF	60	Qy	1 MAQALPWLLWMAGVLPNHGTQHIGIRLRLPRTDEPEEPGRSSF	60	Qy	1 MAQALPWLLWMAGVLPNHGTQHIGIRLRLPRTDEPEEPGRSSF
Qy	61 VEMVDNLRLRKSQGQYYVEMTGVGSPOTUNILVDTGSSNFNAYGAAPHPFLHRHYQROLSS	120	Db	1 MAQALPWLLWMAGVLPNHGTQHIGIRLRLPRTDEPEEPGRSSF	60	Db	1 MAQALPWLLWMAGVLPNHGTQHIGIRLRLPRTDEPEEPGRSSF
Db	61 VEMVDNLRLRKSQGQYYVEMTGVGSPOTUNILVDTGSSNFNAYGAAPHPFLHRHYQROLSS	120	Qy	61 VEMVDNLRLRKSQGQYYVEMTGVGSPOTUNILVDTGSSNFNAYGAAPHPFLHRHYQROLSS	120	Qy	61 VEMVDNLRLRKSQGQYYVEMTGVGSPOTUNILVDTGSSNFNAYGAAPHPFLHRHYQROLSS
Qy	121 YDRLRKGIVYVPTQKWEGELTDLVSIHFGPNVTVRANIAATESDKFEFINGSNNWEGIL	180	Db	121 YDRLRKGIVYVPTQKWEGELTDLVSIHFGPNVTVRANIAATESDKFEFINGSNNWEGIL	180	Qy	121 YDRLRKGIVYVPTQKWEGELTDLVSIHFGPNVTVRANIAATESDKFEFINGSNNWEGIL
Db	121 YDRLRKGIVYVPTQKWEGELTDLVSIHFGPNVTVRANIAATESDKFEFINGSNNWEGIL	145	Db	121 YDRLRKGIVYVPTQKWEGELTDLVSIHFGPNVTVRANIAATESDKFEFINGSNNWEGIL	145	Db	121 YDRLRKGIVYVPTQKWEGELTDLVSIHFGPNVTVRANIAATESDKFEFINGSNNWEGIL
Qy	181 GLAYAEIARPDDSLPLEPFDLSVLKOTHVPMNFSHLCGAGFPQNOSEVLA SVGGSMTIGG	240	Qy	181 GLAYAEIARPDDSLPLEPFDLSVLKOTHVPMNFSHLCGAGFPQNOSEVLA SVGGSMTIGG	240	Qy	181 GLAYAEIARPDDSLPLEPFDLSVLKOTHVPMNFSHLCGAGFPQNOSEVLA SVGGSMTIGG
Db	146 ----- PDDSLPLEPFDLSVLKOTHVPMNFSHLCGAGFPQNOSEVLA SVGGSMTIGG	196	Db	146 ----- PDDSLPLEPFDLSVLKOTHVPMNFSHLCGAGFPQNOSEVLA SVGGSMTIGG	196	Db	146 ----- PDDSLPLEPFDLSVLKOTHVPMNFSHLCGAGFPQNOSEVLA SVGGSMTIGG
Qy	241 DHSLYTGSILWTPIREWWYEVILVRLVEINGDQLKMDCKEYNNDKSIYDGTINRLPK	300	Qy	241 DHSLYTGSILWTPIREWWYEVILVRLVEINGDQLKMDCKEYNNDKSIYDGTINRLPK	300	Qy	241 DHSLYTGSILWTPIREWWYEVILVRLVEINGDQLKMDCKEYNNDKSIYDGTINRLPK
Db	197 DHSLYTGSILWTPIREWWYEVILVRLVEINGDQLKMDCKEYNNDKSIYDGTINRLPK	256	Db	197 DHSLYTGSILWTPIREWWYEVILVRLVEINGDQLKMDCKEYNNDKSIYDGTINRLPK	256	Db	197 DHSLYTGSILWTPIREWWYEVILVRLVEINGDQLKMDCKEYNNDKSIYDGTINRLPK
Qy	301 VFEAAVSKIAASSTEKFDPDFWLGEQLWCWQACTTPWNPFPVTSLYLMGEVTNQSFRIT	360	Qy	301 VFEAAVSKIAASSTEKFDPDFWLGEQLWCWQACTTPWNPFPVTSLYLMGEVTNQSFRIT	360	Qy	301 VFEAAVSKIAASSTEKFDPDFWLGEQLWCWQACTTPWNPFPVTSLYLMGEVTNQSFRIT
Db	257 VFEAAVSKIAASSTEKFDPDFWLGEQLWCWQACTTPWNPFPVTSLYLMGEVTNQSFRIT	316	Db	257 VFEAAVSKIAASSTEKFDPDFWLGEQLWCWQACTTPWNPFPVTSLYLMGEVTNQSFRIT	316	Db	257 VFEAAVSKIAASSTEKFDPDFWLGEQLWCWQACTTPWNPFPVTSLYLMGEVTNQSFRIT
Qy	361 ILPQQYLREVEDATSDQDCYKFAISQSSTGTVMGAVIMEGFYVFDRAKRIGFAYSA	420	Qy	361 ILPQQYLREVEDATSDQDCYKFAISQSSTGTVMGAVIMEGFYVFDRAKRIGFAYSA	420	Qy	361 ILPQQYLREVEDATSDQDCYKFAISQSSTGTVMGAVIMEGFYVFDRAKRIGFAYSA
Db	317 ILPQQYLREVEDATSDQDCYKFAISQSSTGTVMGAVIMEGFYVFDRAKRIGFAYSA	376	Db	317 ILPQQYLREVEDATSDQDCYKFAISQSSTGTVMGAVIMEGFYVFDRAKRIGFAYSA	376	Db	317 ILPQQYLREVEDATSDQDCYKFAISQSSTGTVMGAVIMEGFYVFDRAKRIGFAYSA
Qy	421 HVHDEFRTAVEGPFTLLOMEDCSYNIPOTDESTLMTIAYVMAICALFMLPLCLMVCQW	480	Qy	421 HVHDEFRTAVEGPFTLLOMEDCSYNIPOTDESTLMTIAYVMAICALFMLPLCLMVCQW	480	Qy	421 HVHDEFRTAVEGPFTLLOMEDCSYNIPOTDESTLMTIAYVMAICALFMLPLCLMVCQW
Db	377 HVHDEFRTAVEGPFTLLOMEDCSYNIPOTDESTLMTIAYVMAICALFMLPLCLMVCQW	436	Db	377 HVHDEFRTAVEGPFTLLOMEDCSYNIPOTDESTLMTIAYVMAICALFMLPLCLMVCQW	436	Db	377 HVHDEFRTAVEGPFTLLOMEDCSYNIPOTDESTLMTIAYVMAICALFMLPLCLMVCQW
Qy	481 RCLRCLRQHDDFADDISLK 501		Qy	481 RCLRCLRQHDDFADDISLK 501		Qy	481 RCLRCLRQHDDFADDISLK 501
Db	437 RCLRCLRQHDDFADDISLK 457		Db	437 RCLRCLRQHDDFADDISLK 457		Db	437 RCLRCLRQHDDFADDISLK 432
<b>RESULT 4</b>							
Q9BYB9	PRELIMINARY;	PRT;	432 AA.	O9CUU5	PRELIMINARY;	PRT;	
ID	O9CUU5	ID	266 AA..	AC	O9CUU5;	PRT;	
AC	O9CUU5;	AC		DT	01-JUN-2001 (TREMBrel. 17, Created)		
DT	01-JUN-2001 (TREMBrel. 17, Created)	DT		DT	01-JUN-2001 (TREMBrel. 17, Last sequence update)		
DT	01-JUN-2002 (TREMBrel. 17, Last sequence update)	DT		DE	Beta-site APP cleaving enzyme (Fragment).		
GN	Beta-site APP cleaving enzyme I-432.	GN		DE	BACE.		
OS	Homo sapiens (Human).	OS		GN	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo NCBI_TaxID=9606;	OC		OS	Euteleostomi; Muridae; Murinae; Mus. NCBI_TaxID=10090;		
RN	(1)	RN		RN	[1]		
RP	SEQUENCE FROM N.A.	RP		RP	SEQUENCE FROM N.A.		
RC	TISSUE=BRAIN;	RC		RC	STRAIN=C57BL/6J; TISSUE=BRAIN;		
RX	LINE=21408467; PubMed=11516562;	RX		RX	MEDLINE=11085650; PubMed=11217851;		
RA	Tanahashi H., Tabira T.; Three novel alternatively spliced isoforms of the human beta-site APP cleaving enzyme (BACE) and their effect on amyloid beta-peptide production"; RT	RA		RA	Kawai J., Shinagawa A., Shibusawa K., Yoshino M., Ishii Y., Arakawa T., Hara A., Fukunishi H., Konno H., Adachi J., Kiyosawa H., Kondo S., Yamana T., Aizawa K., Izawa M., Nishi K., Gojobori T., Bono H., Kasukawa T., Saito R., Saito T., Okazaki Y., Matsuda H.A., Ashburner M., Batzov S., Kadota K., Casavant T., Fleischmann W., Gaasterland T., Giissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schiimi L.M., Stabili F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okuda T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Garibaldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Massimi J., Mazzarella J., Mombaerts P., Nordone P., Ringwald M., Rodriguez I., Sakano N., Seya T., Shishiba Y., Storch K.-F., Suzuki H., Sato K., Scheibenbach C., Whittaker C., Wilming L.,		
SQ	PROTEIN: PS00141; ASP-PROTEASE: UNKNOWN 1; SEQUENCE 432 AA.; 48212 MW; 96FC8156FOEED01B CRC64;	SQ		RA			
Query Match	84.8%; Score 2258.5; DB 4; Length 432;	Query Match		RA			





	RESULT	10	Q9P0D2	PRELIMINARY;	PRT;	213 AA.
Qy	ID	Q9P0D2	AC	OP002;		
	DT	01-OCT-2000	(TREMBLrel)	15; Created		
	DT	01-OCT-2000	(TREMBLrel)	15; Last sequence update		
	DT	01-JUN-2002	(TREMBLrel)	21; Last annotation update		
	HSPC104	(Fragment).				
Qy	RA	Zhang Q.H., Ye M., Zhou J., Sheng Y., Wu X.Y., Guan Z.Q., Wang L., Fan H.Y., Mao Y.F., Dai M., Huang Q.H., Chen S.J., Chen Z.;				
Db	RT	"Human partial CDS cloned from cd34+ stem cells."				
	RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.				
	DR	AF11367; AAF28927.; -				
	DR	InterPro: IPR001461; AsproteaseA1.				
	DR	PFam; PF00026; asp; 1.				
	FT	NON_TER 213 AA; 24338 MW;				
	SQ	SEQUENCE FROM N.A.				
	RC	TISSUE-CORD BLOOD;				
	RA	Zhang Q.H., Ye M., Zhou J., Sheng Y., Wu X.Y., Guan Z.Q., Wang L., Fan H.Y., Mao Y.F., Dai M., Huang Q.H., Chen S.J., Chen Z.;				
	RT	"Human partial CDS cloned from cd34+ stem cells."				
	RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.				
	DR	AF11367; AAF28927.; -				
	DR	InterPro: IPR001461; AsproteaseA1.				
	DR	PFam; PF00026; asp; 1.				
	FT	NON_TER 213 AA; 24338 MW;				
	SQ	SEQUENCE FROM N.A.				
Qy	Query Match	26.7%	Score	712.5;	DB	4;
	Best Local Similarity	83.5%	Pred.	2.6e-49;		
Db	Matches	137;	Conservative	4;	Mismatches	12;
					Indels	11;
					Gaps	
Qy	1	GGIDHSLYTGSWLTPIRREWYEVITVREVEINQDLKMDCKEKKYNDKSIVDGTNRL	29			
Db	2	GGIDHSLYTGSWLTPIRREWYEVITVREVEINQDLKMDCKEKKYNDKSIVDGTNRL	60			
Qy	298	PKVYFEAAVYSKIAASSTERFPDFWLGEQLVQAGITPPNIFPVISLYLNGEVTNQSF	35			
	61	PKVYFEAAVYSKIAASSTERFPDFWLGEQLVQAGITPPNIFPVISLYLNGEVTNQSF	12			
358	RITILPQQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEG	401				
Db	121	RITILPQQYLRP-----WKMMWPRPKTTVTVCHLTVING	153			
	RESULT	11	Q9R1P7	PRELIMINARY;	PRT;	255 AA.
	ID	Q9R1P7	AC	OP1P7;		
	DT	01-MAY-2000	(TREMBLrel)	13; Created		
	DT	01-MAY-2000	(TREMBLrel)	13; Last sequence update		
	DT	01-JUN-2002	(TREMBLrel)	21; Last annotation update		
	DE	Aspartyl protease (Fragment).				
	GN	PACE2.				
	OS	Mus musculus (Mouse).				
	OC	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.				
	OX	NCBI_TaxID=10090;				
	RN	SEQUENCE FROM N.A.				
	RP	Accarino M., Fumagalli P., Taramelli R., Ottolenghi S.;				
	RA	"Cloning of a gene from chromosome 21 Down Region encoding a potential transmembrane protease."				

QY	177	EGILGLAYAIFIARPPDSLEPPFDLSLVKQTHVPNLFSLHLCGAGPFLNQSE-VLASVGSSM	235	RA Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M., Cezniker S., Submitted (DEC-2001) to the EMBL/GenBank/NCBI databases.
Db	142	VGILGMGYAVLAKEPDSSITPMVDSIVSVEVTSDFKGLCQ--PLSNSDRELNVQGSKM	198	RL -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
QY	236	TIGGIDHSLVYTG-S-LWTPPIRREWWYEVITVIREINGODLKMDC	278	DR EMBL: AE003630; AAC53016.1;
Db	199	SLGSARPLPSNSDVFVYTAITESFYEVILTNKVGTSDLPC	243	DR HSSP: P06794; 4CMSC
<hr/>				
RESULT 13				
Q9VKP6		PRELIMINARY;	391 AA.	DR FBgap0032304; CGI17134.
ID	Q9VKP6;	PRT;		DR InterPro; IPR001461; Asprotease1.
AC				DR InterPro; IPR01961; Asprotease_site.
DT	01-MAY-2000	(TREMBLrel.	13, Created)	DR Pfam; PF00026; asp_1.
DT	01-TUN-2002	(TREMBLrel.	21, Last annotation update)	DR PRMTS; PR00192; PEPSIN.
DE	CG17134	protein (RE02351P)		DR PROSITE; PS0041; ASP-PROTEASE_2.
GN				DR Aspartyl protease; hydrolase.
RN				DR Matches 105; Conservative 62; Mismatches 145; Indels 49; Gaps 13;
SEQUENCE FROM N.A.				
RC	STRAIN=BERKELEY;			Query Match Score 335.5; DB 5; Length 391;
RX	MEDLINE=20196006; PubMed=10731132;			Best Local Similarity 29.18%; Pred. No. 1.2e-18;
RA	Adams M.D., Celtniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scheer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazquez R.G., Chapple M., Pfleiderer B.D., Wan K.H., Doyle C., Baxter E.G., Heit G., Meklos G.L.G., Abril J.F., Agbayani A., An H.-J., Andrews Pfanckoch C., Baldwin D., Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandaari D., Bolshakov S., Borikova D., Botchan M.R., Bouck J., Brokland P., Brottier P., Burtis K.C., Busam D.A., Butler H., Cedieu E., Center A., Chandra I., Cherry S., Cawley S., Dahlke C., Davenport L., Davies P., De Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Douc L.E., Downes M., Dugan Rocha S., Dunkov B.C., Dunn P., Durbin R., Evangelista C.C., Ferraria S., Fleischmann W., Fosler C., Gabrilian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Hsieg J., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D., Mertulov G., Milashina N.V., Mobarry C., Morris J., Mosheret A., Nelson S.M., Moy M., Murphy B., Muzny D.M., Neilson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reiner K., Remington K., Saunders R.D.C., Scheuler F., Shen H., Shue B.C., Sliden-Klamo I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wasserman R., Waterman S.M., Woodage T., Weinstock G.M., Weissbach J., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; RT "The genome sequence of <i>Drosophila melanogaster</i> ."; RL Science 287:2185-2195(2000).	SEQUENCE FROM N.A.		
RN				RP SEQUENCE FROM N.A.
RC	STRAIN=BERKELEY;			RP STRAIN=PORTO ALEGRE; TISSUE=OVARY;
RA	Champé M., Chávez C., Dorsett V., Dresenek D., Farfan D., Frise E., George R., González M., Guarín H., Kronmiller B., Li P., Liao G., Miranda A., Mongall C.J., Nunoo J., Paragás V., Park S., RT "A home-binding aspartic proteinase from the eggs of the hard tick Boophilus microplus";	AC Q9GX7;	AC Q9GX7;	RC Sorcine M.H.F., Logullo C., Zingali R.B., Baiva-Silva G.O., RA Juliano L., Oliveira P.L.;
RA	Science 287:2185-2195(2000).	DT 01-MAR-2001 (TREMBLrel. 16, Created)	DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)	RA DR J. Biol. Chem. 0:0 (2000);
RA		DT 01-TUN-2002 (TREMBLrel. 21, Last annotation update)	DT 01-TUN-2002 (TREMBLrel. 21, Last annotation update)	RA DR AF286865; ANG0093.1; -.
RA		OS Boophilus microplus (Cattle tick).	OS Eukaryota; Metazoa; Arthropoda; Cheliceraata; Acari; OC Parasitiformes; Ixodida; Ixodidae; Boophilus.	DR HSSP; P00797; 2REN.
RA		OX NCBI_TaxID=6941; [1]	OX MEROPS; A01.054; -.	DR

